

**UNIVERSIDADE FEDERAL DE SANTA CATARINA
CENTRO TECNOLÓGICO
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA
QUÍMICA**

Jarbas Cleber Ferrari

**INFERÊNCIA POR REDES NEURAIS DA DISTRIBUIÇÃO
DE TAMANHO DE PARTÍCULA EM REAÇÕES
CONTÍNUAS DE POLIMERIZAÇÃO EM EMULSÃO**

**Florianópolis
2014**

Jarbas Cleber Ferrari

Inferência por Redes Neurais da Distribuição de Tamanho de Partícula em Reações Contínuas de Polimerização em Emulsão

Tese submetida ao Programa de Pós Graduação em Engenharia Química da Universidade Federal de Santa Catarina para a obtenção do Grau de Doutor em Engenharia Química

Orientador: Prof^ª. Dr^ª. Claudia Sayer
Coorientador: Prof. Dr. Pedro Henrique Hermes de Araújo

Coorientador: Prof^ª. Dr^ª. Fernanda de Castilhos

Florianópolis
2014

Ficha de identificação da obra elaborada pelo autor, através do Programa de Geração Automática da Biblioteca Universitária da UFSC.

Ferrari, Jarbas Cleber

Inferência por Redes Neurais da Distribuição de Tamanho de Partícula em Reações Contínuas de Polimerização em Emulsão / Jarbas Cleber Ferrari, orientadora, Claudia Sayer, coorientadora, Fernanda de Castilhos – Florianópolis, SC, 2013.

161 p.

Tese (doutorado) - Universidade Federal de Santa Catarina, Centro Tecnológico. Programa de Pós-Graduação em Engenharia Química.

Inclui referências

1. Engenharia Química. 2. Distribuição de tamanho de partícula. 3. Função lognormal bimodal. 4. Redes Neurais. I. Sayer, Claudia. II. Universidade Federal de Santa Catarina. Programa de Pós-Graduação em Engenharia Química. Título.

Jarbas Cleber Ferrari

“Inferência por Redes Neurais da Distribuição de Tamanho de Partícula em Reações Contínuas de Polimerização em Emulsão”

Tese de Doutorado julgada adequada para obtenção do Título de “Doutor em Engenharia Química”, e aprovada em sua forma final pelo Programa de Pós Graduação em Engenharia Química da Universidade Federal de Santa Catarina

Florianópolis, 10 de junho de 2014.

Prof.^a. Dr.^a. Claudia Sayer (UFSC)
Orientadora

Prof. Dr. Pedro H. H. Araújo
(UFSC) Coorientador

Prof.^a. Dr.^a. Fernanda de Castilhos
(UFSM) Coorientadora

Prof. Dr. Ricardo A. F. Machado
Coordenador do CPGENQ

Banca Examinadora:

Prof. Dr. Galo Carrilo Le Roux
(USP)

Prof. Dr. Alexandre Ferreira
Santos (UFPR)

Prof.^a. Dr.^a. Cíntia Soares (UFSC)

Prof. Dr. Carlos Alberto Claumann
(UFSC)

Prof. Dr. André Lourenço Nogueira (UFSC)

Este trabalho é dedicado a minha mãe Martina (*in memória*) que foi exemplo de dedicação à família e luta pela vida. Dedico também a meu pai Nolberto que entre muitas coisas me ensinou o valor da integridade e que o verdadeiro amor é feito de doação.

AGRADECIMENTOS

A Deus, que iluminou meu caminho e pacificou meu coração quando as adversidades se multiplicavam.

Aos meus pais Nolberto e Martina, que me mostraram desde cedo o valor do estudo, da honestidade e da família.

A meus dois amores, Ana Carla e Bruno. Vocês trouxeram um significado diferente e especial para minha vida. Esta conquista não seria plena sem o amor e a cumplicidade de vocês.

A minha querida irmã Jane, que sempre foi para mim um porto seguro. Não vou esquecer jamais do seu carinho e generosidade ao me acolher sempre que precisei.

A meu irmão Joel, que do seu jeito sempre soube dizer que estava à disposição para o que der e vier. Sei que sempre contarei com ele e isso me conforta.

A tia Frida, que sempre torceu demais por mim. Tenho certeza que suas preces foram importantes para que eu completasse mais este desafio na vida.

Aos todos aqueles que se consideram meus amigos e que compartilham a alegria desta conquista comigo, especialmente Rogério, Cesar e Edson.

Aos meus orientadores Claudia e Pedro, por sua generosidade, sua experiência e seu profissionalismo.

A orientadora e amiga Fernanda. Sua coragem em enfrentar as securas da vida, sua paciência ao lidar com meu espírito inquieto e sua capacidade de ensinar me inspiram.

Ao programa de Pós-Graduação em Engenharia Química da Universidade Federal de Santa Catarina, pela oportunidade.

*“Os livros não mudam o mundo,
quem muda o mundo são as pessoas.
Os livros só mudam as pessoas”
(Mario Quintana)*

RESUMO

As exigências do mercado mundial em relação à produtividade e qualidade dos produtos poliméricos tornam crescente o número dos trabalhos que tem como objetivo o monitoramento e o controle de qualidade final do látex polimérico. Nesse contexto, a distribuição do tamanho de partícula (DTP) é uma das características mais importantes estando associada às propriedades reológicas, máximo teor de sólidos, adesão e tempo de secagem do látex. No entanto, para distribuições polidispersas ou com partículas muito pequenas, os métodos de determinação experimental das DTPs ao longo da reação de polimerização possuem uma complexidade e requerem um tempo de análise que inviabilizam sua utilização em tempo real. A modelagem das DTPs por meio das Equações de Balanço Populacional, também enfrenta barreiras relacionadas a simplificações que viabilizam a resolução destas equações, mas tornam os resultados limitados. Neste sentido, este trabalho propõe a inferência das curvas de distribuição de tamanho de partícula, através de RNs *Multi Layer Perceptron* (MLP) treinadas pelo algoritmo de otimização global *Particle Swarm Optimization* (PSO). Os dados provenientes de reações de copolimerização em emulsão em reator loop contínuo foram selecionados como objeto de estudo, visto que para as condições reacionais investigadas são observadas sucessivas renucleações que provocam o aparecimento de distribuições multimodais ao longo das reações. O diferencial da estratégia está na implementação de uma rede neural dinâmica em Série (RNDS) e na utilização da Função Lognormal Bimodal (FLB) para descrever as DTPs. Os resultados mostram que a RNDS prediz com precisão a evolução das DTPs ao longo da reação, com alta correlação entre distribuições experimentais e preditas. Além disso, a proposta de utilização de um algoritmo estocástico se mostrou consistente, visto que os valores para os desvios padrão dos parâmetros ajustados se mostraram sempre muito baixos.

Palavras-chave: Distribuição de tamanho de partícula. Função lognormal bimodal. Redes Neurais.

ABSTRACT

Polymer market requirements in terms of productivity and quality have led to an increase of the number of works involving process monitoring and final quality control of polymeric latexes. In this context, the particle size distribution is one of the most important characteristics and being associated with rheological properties, maximum solids content, adhesion and drying time. However, for polydisperse distributions or very small particle sizes, the experimental determination methods of the particle size distributions during the polymerization reaction are quite complex and time consuming hampering their usage in real time monitoring strategies. The particle size distribution modeling by means of Population Balance Equations also faces difficulties related to simplifications, which enable the resolution but turn the results rather limited. Therefore, this work proposes the inference of the particle size distributions by means of a Multi-Layer Perceptron Neural Network trained by the global optimization algorithm Particle Swarm Optimization. As case study, experimental data of emulsion copolymerization reactions in a continuous loop reactor were selected, since the particle size distributions of these reactions present a rather complex and challenging behavior due to successive re-nucleations, which cause the appearance of multimodal distributions during the reactions. The contribution of the suggested strategy is the implementation of a Dynamic Neural Network in series and the usage of the Bimodal Lognormal Function to describe the particle size distribution. The results show that the Dynamic Neural Network predicts with accuracy the evolution of the particle size distributions during the reactions, with high correlation between the experimental and predicted distributions. In addition, the use of a stochastic algorithm has shown consistency, since the values to the standard deviation of the adjusted parameters were very low.

Keywords: Particle Size Distributions. Bimodal Lognormal Function. Neural Networks.

LISTA DE FIGURAS

Figura 1 Mecanismo cinético de nucleação e terminação dos radicais poliméricos em uma polimerização em emulsão.....	38
Figura 2 Representação do mecanismo de nucleação micelar	38
Figura 3 Representação do mecanismo de nucleação homogênea	39
Figura 4 Interações ocorridas durante o mecanismo de polimerização e as relações com a DTP	41
Figura 5 Estrutura elementar de uma rede neural.....	51
Figura 6 Estrutura básica de uma Rede Neural Multi-layer Perceptron	53
Figura 7 Representação vetorial das influências sobre o movimento das partículas X_i	61
Figura 8 Fluxograma de implementação do algoritmo PSO	63
Figura 9 Dados de distribuição (a) em escala linear e (b) em escala logarítmica.....	66
Figura 10 Esquema do reator loop	72
Figura 11 Sistema de alimentação da pré-emulsão	72
Figura 12 Fluxograma de implementação da Rede Neural Dinâmica....	79
Figura 13 Fluxograma identificando variáveis de entrada e saída da rede dinâmica em série para uma função lognormal bimodal	85
Figura 14 Dados experimentais (ARAÚJO, 1999) e calculados pela RND para a conversão na reação R5.....	88
Figura 15 Dados experimentais (ARAÚJO,1999) e calculados pela RND para a conversão na reação R9.....	89

Figura 16 Comparação entre dados experimentais (ARAÚJO, 1999) e calculados pela rede neural MLP para o diâmetro médio de partículas na reação R5	90
Figura 17 Comparação entre dados experimentais (ARAÚJO, 1999) e calculados pela rede neural MLP para o diâmetro médio de partículas na reação R9	91
Figura 18 Dados experimentais (ARAÚJO, 1999) e calculados pela RND com dados atrasados de Dp médio na inferência do Dp médio na reação R5.....	92
Figura 19 Dados experimentais (ARAÚJO, 1999) e calculados pela RND com dados atrasados de Dp médio na inferência do Dp médio na reação R9.....	92
Figura 20 Curvas de DTPs experimentais e preditas pela FLT	94
Figura 21 Curvas de DTPs experimentais e preditas pela FLB.....	97
Figura 22 Parâmetro M1 ajustado da FLB e predito pela RNDS.....	102
Figura 23 Parâmetro M2 ajustado da FLB e predito pela RNDS.....	102
Figura 24 Parâmetro D1 ajustado da FLB e predito pela RNDS	103
Figura 25 Parâmetro D2 ajustado da FLB e predito pela RNDS	103
Figura 26 Parâmetro P1 ajustado da FLB e predito pela RNDS	104
Figura 27 Parâmetro P2 ajustado da FLB e predito pela RNDS	104
Figura 28 Comparação entre curvas de DTPs experimentais e preditas da reação R5	107
Figura 29 Efeito da temperatura na conversão (●) T = 60° C [R2], (▲) T = 50° C [R4]	139
Figura 30 Efeito da temperatura no diâmetro médio de partícula (●) T = 60° C [R2], (▲) T = 50° C [R4]	140

Figura 31 Efeito da concentração de emulsificante na evolução da conversão (●) na reação R7.....	140
Figura 32 Efeito da concentração de emulsificante na evolução do diâmetro médio de partícula (●) na reação R7	141
Figura 33 Efeito da concentração de iniciador sobre a conversão. (●) 0,30% de iniciador (R2) e (▲) 0,18% de iniciador (R8).....	141
Figura 34 Efeito da concentração de iniciador sobre o diâmetro médio. (●) 0,30% de iniciador (R2) e (▲) 0,18% de iniciador (R8)	142
Figura 35 Velocidade de crescimento das diferentes populações de partícula em (▲) T = 60oC (R2) e (●) T = 50oC (R4)	144
Figura 36 Evolução das distribuições de tamanho de partícula entre as amostras 15 e 19 na reação R2.....	147
Figura 37 Superfícies geradas a partir das DTPs, para as reações realizadas por Araújo (1999).....	149
Figura 38 Resultados da função objetivo [NSSE] para o treinamento e validação da rede neural MLP com diferentes arquiteturas.....	155
Figura 39 Resultados da função objetivo [NSSE] para treinamento e validação da rede neural com diferentes números de neurônios na camada oculta na estimação do diâmetro médio de partícula.	157
Figura 40 Resultados do treinamento e validação da rede neural com diferentes números de neurônios na camada oculta na estimação da conversão.....	158
Figura 41 Resultados do treinamento e validação da rede neural com diferentes números de neurônios na camada oculta na estimação do parâmetro M1	159
Figura 42 Resultados do treinamento e validação da rede neural com diferentes números de neurônios na camada oculta na estimação do parâmetro D2.....	160

LISTA DE TABELAS

Tabela 1 Mecanismo cinético observado na polimerização em emulsão	37
Tabela 2 Modelos de balanço populacional em polimerização em emulsão	48
Tabela 3 Trabalhos que contribuíram para a evolução do PSO	62
Tabela 4 Condições de operação das Reações de Polimerização realizadas por Araújo (1999).....	74
Tabela 5 Intervalo de busca dos parâmetros da FLM.....	81
Tabela 6 Codificação dos parâmetros da função lognormal bimodal.....	84
Tabela 7 Erro percentual médio obtido nesse trabalho e por Araújo <i>et. al</i> (2001b), na predição da conversão para o grupo de validação.....	87
Tabela 8 Coeficiente de correlação entre dados experimentais e preditos para a conversão pela RND	89
Tabela 9 Erro percentual médio obtido nesse trabalho e por Araújo <i>et. al</i> (2001b), na predição do Dp médio para o grupo de validação.....	90
Tabela 10 Coeficiente de correlação entre dados experimentais e preditos para o Dp médio pela RND sem e com dados atrasados de Dp médio.....	91
Tabela 11 Parâmetros das curvas de DTP estimados usando FLT.....	95
Tabela 12 Parâmetros das curvas de DTP estimados usando FLB.....	99
Tabela 13 Erros médios percentuais da RND e da RNDS	101
Tabela 14 Parâmetros e desvios padrão da FLB obtidos na inferência por RNDS.....	105
Tabela 15 Coeficientes de correlação entre DTPs experimentais típicas e inferidas pela RNDS	110

Tabela 16 Dados da distribuição de tamanho de partícula das amostras 15 e 19 na reação R2	146
Tabela 17 Matriz de Correlação entre os parâmetros ajustados da Função Lognormal Trimodal [Reação – R5]	151
Tabela 18 Matriz de Correlação entre os parâmetros ajustados da Função Lognormal Trimodal [Reação – R8]	151
Tabela 19 Matriz de Correlação entre os parâmetros ajustados da Função Lognormal Bimodal [Reação – R5].	152
Tabela 20 Matriz de Correlação entre os parâmetros ajustados da Função Lognormal Bimodal [Reação – R4]	152
Tabela 21 Matriz de Correlação entre os parâmetros ajustados da Função Lognormal Bimodal [Reação – R8]	152
Tabela 22 Valores de crítico para o Teste de Kolmogorov-Smirnov com diferentes níveis de significância (α).....	154
Tabela 23 Resultado da NSSE na inferência do diâmetro médio de partícula com diferentes arquiteturas para RNs MLP com duas camadas ocultas.....	156
Tabela 24 Resultado da NSSE e desvio padrão na inferência do diâmetro médio de partícula com diferentes números de neurônios na camada oculta.....	157
Tabela 25 NSSE e desvio padrão na inferência da conversão com diferentes números de neurônios na camada oculta	158
Tabela 26 NSSE e desvio padrão na inferência do parâmetro M1, com diferentes números de neurônios na camada oculta	160
Tabela 27 NSSE e desvio padrão na inferência do parâmetro D2, com diferentes números de neurônios na camada oculta	161

SUMÁRIO

1 INTRODUÇÃO	33
1.1 OBJETIVO GERAL.....	34
1.2 OBJETIVOS ESPECÍFICOS.....	34
2 REVISÃO BIBLIOGRÁFICA.....	35
2.1 POLIMERIZAÇÃO EM EMULSÃO.....	35
2.2 A DISTRIBUIÇÃO DO TAMANHO DE PARTÍCULAS NA REAÇÃO DE POLIMERIZAÇÃO	40
2.3 MODELOS MATEMÁTICOS DA DISTRIBUIÇÃO DO TAMANHO DE PARTICULAS EM POLIMERIZAÇÃO	45
2.4 REDES NEURAIIS ARTIFICIAIS	50
2.4.1 Princípios.....	50
2.4.2 Redes Neurais Multi-Layer Perceptron	52
2.4.3 Aplicações de Redes Neurais ao Processo de Polimerização em Emulsão	56
2.5 PARTICLE SWARM OPTIMIZATION	59
2.6 DISTRIBUIÇÕES DE PROBABILIDADE	63
2.6.1 Princípios.....	63
2.6.2 Distribuição Lognormal	65
2.6.3 Aplicações da Função Densidade Lognormal.....	68
2.7 CONSIDERAÇÕES SOBRE A REVISÃO BIBLIOGRÁFICA	70
3. MATERIAL E MÉTODOS.....	71
3.1 DADOS EXPERIMENTAIS.....	71
3.2 IMPLEMENTAÇÃO DO PSO.....	75
3.3 DETERMINAÇÃO DA ARQUITETURA DA REDE NEURAL <i>MULTI-LAYER PERCEPTRON</i>	76

3.4 DESENVOLVIMENTO DO MODELO NEURAL PARA PREDIÇÃO DE CONVERSÃO E DIÂMETRO MÉDIO DE PARTÍCULA	77
3.5 MODELAGEM DA DISTRIBUIÇÃO DE TAMANHO DE PARTÍCULA USANDO FUNÇÃO LOGNORMAL MULTIMODAL	80
3.6 DESENVOLVIMENTO DO MODELO NEURAL HÍBRIDO PARA INFERÊNCIA DAS DTPs.....	82

4. RESULTADOS E DISCUSSÕES 87

4.1 INFERÊNCIA DA CONVERSÃO E DIÂMETRO MÉDIO	87
4.1.1 Arquitetura da Rede Neural MLP	87
4.1.2 Inferência da Conversão	87
4.1.3 Inferência do Diâmetro Médio de Partícula.....	89
4.2 MODELAGEM DAS DISTRIBUIÇÕES DE TAMANHO DE PARTÍCULA USANDO A FUNÇÃO LOGNORMAL MULTIMODAL	93
4.3 INFERÊNCIA DAS DISTRIBUIÇÕES DE TAMANHO DE PARTÍCULA POR REDES NEURAIS DINÂMICAS EM SÉRIE	100
4.3.1 Inferência dos Parâmetros da Função lognormal Bimodal	100
4.3.2 Construção das DTPs preditas.....	107

5 CONCLUSÕES E SUGESTÕES.....113

5.1 CONCLUSÕES	113
5.2 SUGESTÕES	114

REFERÊNCIAS.....117

APÊNDICE A – Modelagem da Distribuição de Tamanho de Partícula fazendo uso das Equações de Balanço Populacional135

APÊNDICE B – Efeitos das condições de operação sobre a conversão e diâmetro médio de partícula.....139

APÊNDICE C – Interpolações lineares da curvas de distribuições de tamanho de partícula.....	143
APÊNDICE D – Comportamento da distribuição de tamanho de partícula ao longo das reações.....	149
APÊNDICE E – Matriz de correlação entre parâmetros ajustados da Função Lognormal Bimodal.....	151
APÊNDICE F – Teste de Kolmogorov-Smirnov.....	153
APÊNDICE G – Determinação da Arquitetura da Rede Neural MLP	155

LISTA DE ABREVIATURAS E SIGLAS

DTP	Distribuição de Tamanho de Partícula
RN	Rede Neural
<i>MLP</i>	<i>Multi Layer Perceptron</i>
<i>PSO</i>	<i>Particle Swarm Optimization</i>
RND	Rede Neural Dinâmica
RNDS	Rede Neural Dinâmica em Série
FLB	Função Lognormal Bimodal
<i>PSD</i>	<i>Particle Size Distribution</i>
<i>DNN</i>	<i>Dynamic Neural Network</i>
<i>BLF</i>	<i>Bimodal Lognormal Function</i>
<i>SDNN</i>	<i>Serial Dynamics Neural Network</i>
<i>DLS</i>	<i>Dynamic Light Scattering</i>
<i>ELS</i>	<i>Elastic Light Scattering</i>
Dp	Diâmetro da partícula
RND	Rede Neural Dinâmica
CMC	Concentração Micelar Crítica
<i>QELS</i>	<i>Quasi-Elastic Light Scattering</i>
<i>TEM</i>	<i>Transmission Electron Microscopy</i>
<i>CHFD</i>	<i>Capillary Hydrodynamic Fractionation</i>
EBP	Equações do Balanço Populacional
<i>CFD</i>	<i>Computational Fluid Dynamics</i>
GA	<i>Genetic Algorithm</i>
NSSE	Soma dos Erros Quadráticos Normalizados
FLM	Função Lognormal Multimodal
FLT	Função Lognormal Trimodal
SSE	Soma dos Erros Quadráticos

LISTA DE SÍMBOLOS

I	Iniciador
$R\bullet$	Radicais Livres
M	Monômero
$P_1\bullet$	Radical Primário
$P_n\bullet$	Cadeia Polimérica
r	raio da partícula
$F(r,t)$	função densidade populacional das partículas
dr/dt	taxa de crescimento das partículas de raio r
\mathcal{R}_{nuc}	taxa de formação de novas partículas
\mathcal{R}_{coal}	taxa de coalescência das partículas.
\mathcal{R}_{form}	taxa de formação das partículas
\mathcal{R}_{cons}	taxa de consumo das partículas
$B(r,r')$	coeficiente da taxa de coalescência de duas partículas de raios r e r'
V_{aq}	volume da fase aquosa.
$\mathcal{R}_{micelar}$	taxa de nucleação micelar
$\mathcal{R}_{homogênea}$	taxa de nucleação homogênea
z	tamanho mínimo da nova partícula
$k_{emic,i}$	coeficiente da taxa de captura de radicais pelas micelas
$[R_m^*]_{aq}$	concentração de radicais na fase aquosa de tamanho m
j_{crit}	tamanho crítico da partícula
P_{wi}	probabilidade de que o radical na fase aquosa seja terminado por um mero do tipo i
$k_{p,av}^{aq}$	coeficiente da taxa de propagação na fase aquosa
$[R_{j_{crit}-1}^*]_{aq}$	concentração de radicais de tamanho $j_{crit} - 1$ na fase aquosa.
n_A	números de radicais tipo A
n_B	números de radicais tipo B
$\bar{n}(r,t)$	número médio de radicais por partícula em função do tamanho das partículas
$k_{e,i}^m$	coeficiente da taxa de captura de radicais de tipo i e tamanho m

k_{dMj}	coeficiente da taxa de saída de radicais
D_{wj}	coeficiente de difusão do monômero j na fase aquosa
k_{pil}	coeficiente da taxa de propagação de um radical terminado em j com monômero do tipo l
$V_p(r)$	volume de uma partícula de tamanho r
k_{trij}	coeficiente da taxa de transferência de cadeia de um radical terminado em i para um monômero do tipo j
k_{tav}	coeficiente da taxa de terminação para pseudo-homopolimerização
$[M_j]_p$	concentração do monômero j nas partículas
k_{pij}	coeficiente da taxa de propagação de um radical terminado em i com monômero do tipo j
N_A	número de Avogadro
MM_j	massa molecular do monômero j
p_i	probabilidade que um radical nas partículas seja de tipo i
y_k	saída do k -ésimo neurônios da camada interna
Z_p	saída do p -ésimo neurônio da camada de saída da rede neural
b_k	— bis ” do k -ésimo neurônio da camada interna
b_p	— bis ” do p -ésimo neurônio da camada de saída
w^h	peso da conexão do i -ésimo neurônio da camada de entrada para o k -ésimo neurônio da camada interna
w_{kp}^o	peso da conexão do k -ésimo neurônio da camada interna para o p -ésimo neurônio da camada de saída
n	número de entradas
m	número de unidades internas
S	espaço de busca
R^+	números reais positivos
X_i^k	vetor posição da partícula i no espaço k -dimensional
V_i^k	vetor velocidade da partícula i no espaço k -dimensional
$x_n^{\text{máx}}$	limite superior do vetor posição n

x_n^{\min}	limite inferior do vetor posição n
V_n^{\max}	velocidade máxima da partícula n n = número total de partículas
r_n	número randômico da coordenada n
P_i^k	melhor posição encontrada pela partícula i até a k -ésima iteração
P_g^k	melhor posição encontrada por todo o grupo até a k -ésima iteração
w	Peso inercial
c_1	Coefficiente de aceleração
c_2	Coefficiente de aceleração
$f(x)$	função densidade de probabilidade
X	variável aleatória contínua
S	área gerada a partir da função densidade de probabilidade
$g(x)$	função densidade de probabilidade unitária
$Y = \log(X)$	variável aleatória contínua em escala logarítmica
f_i	frequências absolutas do agrupamento i
x_i	
N	somatória das frequências absolutas
w^i	fração da moda i
f_i^n	distribuição de frequência normalizada do agrupamento i
w_o	peso inercial inicial
u	parâmetro de controle do PSO
P_{best}	melhor resultado encontrado
P_n	n últimos melhores resultados
N_{pesos}	número de pesos da rede neural
e	número de entradas da rede neural
c	número de neurônios na camada oculta de rede neural
s	número de saídas da rede neural
h	fração do <i>holdup</i>
$[i]$	concentração do iniciador
$[s]$	concentração do surfactante
$[X]$	Conversão

h	fração do <i>holdup</i>
t_k^n	saída normalizada desejada para a unidade de saída k para o padrão n
o_k^n	saída normalizada obtida pela rede na unidade k da camada de saída para o padrão n
x_{exp}^i	valor experimental para o ponto experimental
x_{cal}^i	valor estimado pela rede neural para o ponto experimental i
d_k	diâmetro da partícula no ponto k .
f_{exp}^k	valor experimental da frequência no ponto k
M1	média geométrica da moda 1
M2	média geométrica da moda 2
D1	desvio padrão geométrico da moda 1
D2	desvio padrão geométrico da moda 2
P1	fração da moda 1
P2	fração da moda 2

Letras gregas

Λ_{n+m}	terminação por combinação
$\Lambda_n + \Lambda_m$	terminação por Desproporcionamento
ρ_{pol}	densidade do polímero
μ_g^i	média geométrica da moda i
σ_g^i	desvio padrão geométrico da moda i
ρ	Coefficiente de correlação

1 INTRODUÇÃO

A polimerização em emulsão tem sido amplamente utilizada em uma variedade de processos para a produção de látex sintéticos, isso porque apresenta vantagens como a baixa viscosidade, a possibilidade do controle da massa molecular, o uso da água como meio contínuo, além da possibilidade de se obter conversões altas (KIPARISSIDES, 1996; CHERN, 2006).

Em um processo de polimerização em emulsão, o intuito é a obtenção de látex polimérico com propriedades mecânicas e estruturais específicas para determinadas aplicações. Por isso, muitos trabalhos vêm usando estimadores *on-line* para controlar variáveis como a viscosidade, conversão e massa molecular ou para determinar condições iniciais ideais e perfis de temperatura em estratégias de otimização e controle (FERNANDES e LONA, 2005).

Segundo Vale e McKenna, 2005, a distribuição do tamanho de partícula (DTP) é uma das características mais importantes em relação à qualidade do látex polimérico, determinando suas propriedades reológicas, máximo teor de sólidos, adesão e tempo de secagem.

No entanto, a determinação experimental das curvas de distribuição de tamanho de partícula fica prejudicada quando as distribuições são polidispersas e ou as partículas poliméricas são muito pequenas (≤ 100 nm), acrescentando complexidade ao procedimento e aumentando consideravelmente o tempo de análise, inviabilizando sua aplicação em tempo real.

A modelagem das DTPs em reações de polimerização em emulsão, considerando as equações de balanço populacional também possui restrições quanto à dependência de um número grande de parâmetros, frequentemente desconhecidos, e ao alto custo computacional para resolver modelos cinéticos e ou hidrodinâmicos completos.

Uma alternativa para o problema pode ser a determinação inferencial das curvas de distribuição de tamanho de partícula por RNs (RNs). Para tanto, algumas limitações devem ser superadas quanto à consistência do conjunto de dados experimentais, precisão da função que descreve as DTPs e robustez da estratégia de treinamento da rede neural.

A literatura científica não tem explorado esta alternativa se apresentando, portanto, como um desafio. A falta de quantidade e diversidade de distribuições experimentais constitui provavelmente o fator limitante. Proposta similar é apresentada por Gugliotta *et. al*

(2009), no entanto, as DTPs são decorrentes de reações simuladas de poliestireno e a rede neural tem como entrada medidas indiretas do tamanho das partículas obtidas por DLS (*Dynamic Light Scattering*) e ELS (*Elastic Light Scattering*).

1.1 OBJETIVO GERAL

Realizar a inferência por RNs das DTPs de reações de polimerização em emulsão em reator tubular loop contínuo.

1.2 OBJETIVOS ESPECÍFICOS

- ✓ Investigar a consistência matemática da função lognormal em descrever distribuições de tamanho de partícula polidispersas.

- ✓ Avaliar a eficiência da rede neural dinâmica quando treinada pelo algoritmo de otimização global PSO.

- ✓ Realizar a inferência das DTPs ao longo das reações fazendo uso de variáveis de operação como entradas da rede neural dinâmica.

Esta Tese esta organizada da seguinte forma: O cap.02 apresenta uma revisão da literatura considerando aspectos referentes às reações de polimerização em emulsão, a importância da distribuição de tamanho de partícula, a modelagem das DTPs, a utilização de NR no desenvolvimento de *soft-sensors* para o monitoramento e controle de reações de polimerização e aos princípios e aplicações da função distribuição lognormal, O cap.03 discute as características essenciais do conjunto de dados experimentais, a implementação do PSO, a arquitetura da rede neural dinâmica para inferência da conversão e diâmetro médio de partícula (D_p médio), a implementação da função lognormal multimodal e finalmente a arquitetura da rede neural dinâmica em série (RNDS) para inferência das DTPs; O cap. 04 apresenta e discute os resultados da inferência por RND da conversão e D_p médio, da escolha pela função lognormal bimodal na construção das DTPs e da inferência das DTPs por RNDS; O cap. 05 exhibe conclusões sobre o trabalho, sugerindo caminhos a serem seguidos em trabalhos correlatos futuros.

2 REVISÃO BIBLIOGRÁFICA

Este capítulo tem como objetivo inicial apresentar a técnica e o mecanismo da polimerização em emulsão, bem como a influência das variáveis de operação sobre a determinação das curvas de distribuição de tamanho de partícula, justificando a relevância destas curvas na qualidade final do látex polimérico. Busca também realizar uma discussão sobre a modelagem matemática da distribuição do tamanho de partícula, incluindo as equações do balanço populacional, explorando suas limitações e abordagens de resolução. Discute aspectos da implementação de RNs, em particular das RNs *Multi Layer Perceptron*, no desenvolvimento de *soft-sensors* para o monitoramento e controle de reações de polimerização em emulsão, enfatizando as principais aplicações das RNs na inferência de variáveis importantes na condução das reações. Apresenta ainda os princípios e aplicações do algoritmo PSO como método estocástico de estimação de parâmetros. Finaliza mostrando os fundamentos matemáticos e aplicações da função lognormal na modelagem de distribuições mono ou polidispersas em diversas áreas de conhecimento.

2.1 POLIMERIZAÇÃO EM EMULSÃO

A polimerização em emulsão é caracterizada por ocorrer em sistemas heterogêneos, onde a maior parte da reação de propagação da cadeia ocorre nas partículas poliméricas (20 a 1000 nm de diâmetro) dispersas na fase aquosa (ARAÚJO, 1999).

Os principais componentes da polimerização em emulsão incluem, além da água e do monômero, o iniciador, o emulsificante e agentes de transferência de cadeia. Especificamente, observa-se que a água é o meio dispersante formando a fase contínua e agindo de modo a manter a viscosidade baixa. Além disso, propicia uma boa transferência de calor e promove a transferência de monômero das gotas para as micelas ou partículas poliméricas. A fase contínua também é local de decomposição do iniciador e formação dos oligômeros e meio para transferência dinâmica de surfactante entre as fases. O emulsificante (também conhecido como surfactante ou tensoativo) consiste em uma longa cadeia hidrocarbonatada de natureza hidrofóbica com uma extremidade hidrofílica, tendo função de estabilizar a reação. O iniciador, usualmente hidrossolúvel, quando adicionado ao reator sofre decomposição (ativada por via térmica, química ou fotoquímica), tendo o papel de reagir com o monômero e formar um centro ativo iniciando a geração de radicais a uma velocidade que depende da natureza química do iniciador, do pH do meio e da temperatura do sistema. Os agentes de

transferência de cadeia podem ser adicionados ao meio reacional com a função de controlar a massa molecular do polímero formado (GILBERT, 1995).

Com uma produção mundial anual que excede 20 milhões de toneladas, a polimerização em emulsão tem sido amplamente utilizada em uma variedade de processos para a produção de látex sintéticos desde sua introdução em escala industrial, em função de uma necessidade gerada pela Segunda Guerra Mundial (WARSON e FINCH, 2001). O látex polimérico sintético, preparado pelo processo de polimerização em emulsão, é usado como *commodities* em uma grande variedade de aplicações, tais como: borrachas sintéticas, adesivos, tintas, aditivos para papel, tecidos e material de construção, polímeros de alto impacto, testes diagnósticos e sistema de liberação de fármacos (MEYER e KEURENTJES, 2005).

Uma das principais vantagens deste processo está na possibilidade de obter simultaneamente polímeros de alta massa molecular e elevadas velocidades de polimerização (CHERN, 2006). Outras vantagens são: A baixa viscosidade do látex permite maior dissipação de calor, o que facilita o controle de temperatura; A possibilidade do controle da massa molecular pela adição de agentes de transferência de cadeia; A utilização de água como meio contínuo, minimizando os efeitos tóxicos e aumentando a segurança da operação; A possibilidade de produção de partículas estruturadas do tipo casca-núcleo, alterando as características morfológicas do produto final e aumentando as possibilidades de aplicação; A possibilidade de ser realizada com altas conversões, aumentando sua eficiência e diminuindo problemas com monômero residual (KIPARISSIDES, 1996).

Existem dois mecanismos básicos de formação de cadeias poliméricas: a polimerização em cadeia e a polimerização em etapas. O mecanismo de formação das cadeias poliméricas na emulsão é a polimerização em cadeia, e nesse caso a reação se inicia a partir de uma espécie reativa produzida pelo iniciador ou catalisador. A espécie ativa reage com a molécula de monômero pelo rompimento de ligações π , formando uma nova espécie ativa e o crescimento da cadeia se dá pela adição de unidades monoméricas de forma repetitiva, sempre a partir do centro ativo, em tempos relativamente curtos.

A polimerização em cadeia é caracterizada pelas reações apresentadas na Tabela 1 e envolve três etapas principais e consecutivas no tempo de vida de cada cadeia polimérica. A etapa de iniciação consiste na decomposição do iniciador (I) formando um par de radicais livres ($R\bullet$) e na adição de um destes radicais a uma molécula de

monômero (M), produzindo um radical primário ($P_1\bullet$). Na etapa de propagação há o crescimento da espécie iniciadora de cadeia ($P_1\bullet$) através de adições sucessivas de moléculas de monômero. O crescimento da cadeia polimérica ($P_n\bullet$) é interrompido pela reação bimolecular de radicais ou de desativação monomolecular na etapa de terminação. A terminação pode ocorrer por combinação, formando uma molécula de polímero (Λ_{n+m}), ou por desproporcionamento, formando duas moléculas de polímero ($\Lambda_n + \Lambda_m$). Adicionalmente, dependendo do monômero utilizado, podem ocorrer reações de transferência de cadeia, em que ocorre a terminação de uma cadeia polimérica em crescimento ($P_n\bullet$) deslocando o radical para uma molécula de monômero (M) ou uma cadeia polimérica formada (Λ_m) (GILBERT, 1995; ODIAN, 2004).

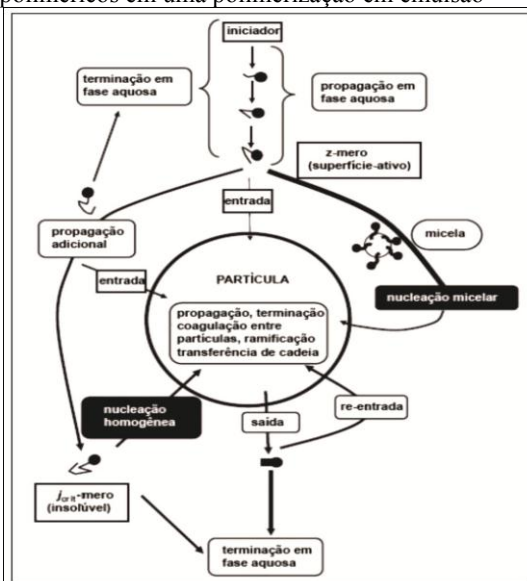
Tabela 1 Mecanismo cinético observado na polimerização em emulsão

<i>Iniciação</i>	$I \xrightarrow{k_i} 2R\bullet$
	$R\bullet + M \xrightarrow{k_{im}} P_1\bullet$
Propagação	$P_n\bullet + M \xrightarrow{k_p} P_{n+1}\bullet$
Transferência de cadeia p/ monômero	$P_n\bullet + M \xrightarrow{k_{tm}} \Lambda_n + P_1\bullet$
Transferência de cadeia p/ polímero	$P_n\bullet + \Lambda_m \xrightarrow{k_{tp}} \Lambda_n + P_m\bullet$
Terminação por combinação	$P_n\bullet + P_m\bullet \xrightarrow{k_{tc}} \Lambda_{n+m}$
Terminação por desproporcionamento	$P_n\bullet + P_m\bullet \xrightarrow{k_{td}} \Lambda_n + \Lambda_m$

Fonte: Gilbert (2005)

Apesar das vantagens, a polimerização em emulsão é um processo bastante complexo porque a nucleação, o crescimento e a estabilidade de partículas poliméricas são etapas controladas pelos mecanismos de polimerização via radicais livres em combinação com vários fenômenos coloidais (CHERN, 2006). A Figura (1) apresenta em síntese os mecanismos envolvidos em uma polimerização em emulsão.

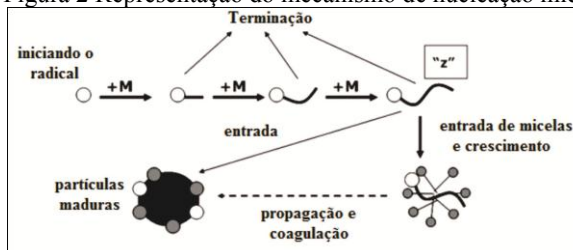
Figura 1 Mecanismo cinético de nucleação e terminação dos radicais poliméricos em uma polimerização em emulsão



Fonte: Adaptado de Coen *et. al* (2004)

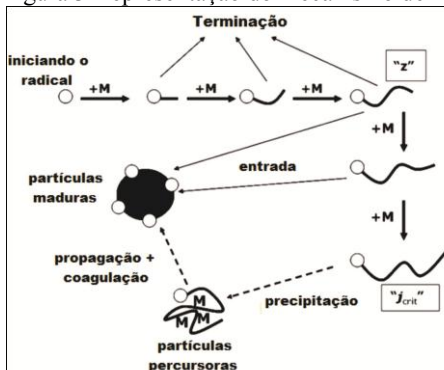
Nas reações de polimerização em emulsão as partículas de polímero podem ser formadas pela entrada de radicais oligoméricos nas micelas (agregados coloidais que se formam quando a concentração de surfactante, na fase aquosa, é superior a concentração micelar crítica CMC), sendo neste caso denominada nucleação micelar (Figura 2) Ou ainda, estes radicais poliméricos continuam a propagar com os monômeros dissolvidos na fase aquosa até um tamanho crítico, formando novas partículas poliméricas por precipitação, sendo denominada nucleação homogênea (Figura 3).

Figura 2 Representação do mecanismo de nucleação micelar



Fonte: Adaptado de Thickett e Gilbert (2007)

Figura 3 Representação do mecanismo de nucleação homogênea



Fonte: Adaptado de Thickett e Gilbert (2007)

Embora o período de nucleação seja bastante curto, a formação de partículas durante a fase inicial da polimerização tem um papel crucial, determinando o tamanho de partícula de látex final e distribuição de tamanho de partícula (THOMSON, YASZEMSKI e POWERS, 1995).

A probabilidade de o radical entrar nas gotas de monômeros (nucleação das gotas) é baixa devido à área superficial ser muito pequena quando comparado a área superficial das micelas. Assim, as partículas são formadas, preferencialmente, por nucleação micelar. Nesse caso, após o período de nucleação das partículas poliméricas, ocorre o crescimento destas. No decorrer da reação, o surfactante presente na fase aquosa migra para a superfície das partículas a fim de estabilizar estas. Esta migração ocorre até que as micelas se esgotem, indicando o final do período de nucleação micelar. Na sequência a reação continua, caracterizando a fase de crescimento das partículas, até o consumo do monômero presente nas gotas de monômero, resultando no seu desaparecimento. Finalmente, a reação prossegue até o consumo completo do monômero presente nas partículas de polímero. Esta etapa é caracterizada pela redução da concentração de monômero nas partículas, ocorrendo um aumento da viscosidade das partículas (efeito gel), gerando um aumento do número médio de radicais por partícula resultando no aumento da taxa de polimerização (MEYER e KEURENTJES, 2005; CHERN, 2006; THICKETT e GILBERT, 2007).

Segundo Chern (2006), os fatores fundamentais que governam a fase de crescimento das partículas são: o transporte de monômero, radicais livres e surfactante para as partículas crescentes e a partição destes reagentes entre a fase contínua, as gotas de monômero

emulsionadas, as partículas de polímero inchadas pelo monômero e a interface óleo-água.

Muito em função das suas inúmeras vantagens e aplicações, as condições de processo de polimerização em emulsão que definem as características do látex polimérico formado ainda são investigadas por institutos de pesquisa e indústrias. O estado da arte do mecanismo e cinética da polimerização em emulsão pode ser encontrado com detalhes em Meyer e Keurentjes (2005), Chern (2006), Thickett e Gilbert (2007).

2.2 A DISTRIBUIÇÃO DO TAMANHO DE PARTÍCULAS NA REAÇÃO DE POLIMERIZAÇÃO

Para sistemas de polimerização, as características de composição fina (microestrutura das cadeias poliméricas), de morfologia (geralmente não mensuráveis) e coloidais (diâmetro e distribuição de diâmetros de partículas), determinam as propriedades físicas, químicas, reológicas e mecânicas, e como consequência, as propriedades de uso dos produtos obtidos (GIORDANI, 2003).

Nesse contexto, DTP é uma das características mais importantes do látex, determinando suas propriedades reológicas, conteúdo máximo de sólidos, adesão e tempo de secagem (VALE e McKENNA, 2005). De forma prática, na produção de emulsões de tintas e vernizes, a distribuição do tamanho das partículas de polímero tem uma influência crítica na qualidade do espalhamento do látex sobre as superfícies recobertas, na capacidade de cobertura, penetração, formação de filme e estabilidade, já que, nestes casos, o látex é o produto direto da polimerização (ARAÚJO, 1999).

A distribuição de tamanhos de partículas pode, ao longo do processo de polimerização, passar por variações significativas não apenas do tamanho médio, mas também da forma de distribuição (unimodal ou multimodal, estreita ou larga), sendo que os mecanismos envolvidos nestas variações usualmente estão relacionados a um ou mais dos seguintes aspectos: a cinética de reação, a termodinâmica do processo, a agitação e os fenômenos de microescala, incluindo a transferência de massa e de calor entre diferentes fases presentes no sistema (ULIANA, 2007).

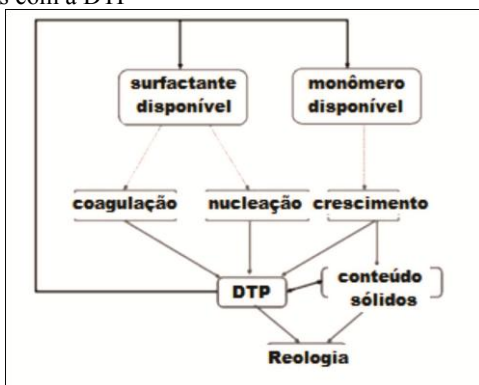
De acordo com Thickett e Gilbert (2007) a fase de nucleação das partículas controla o tamanho médio e a distribuição dos tamanhos das partículas do produto final. O mecanismo de nucleação determina o número de partículas em uma polimerização em cadeia, sendo este uma função da concentração de iniciador e surfactante, de forma que se não

houver nucleações secundárias, o número se mantém constante durante toda a reação.

Informações relativas ao mecanismo de formação de partícula podem ser obtidas da distribuição de tamanho de partícula durante a nucleação. Porém, a interpretação é dificultada devido à falta de exatidão dos equipamentos existentes para determinar DTP quando partículas pequenas e polidispersas fazem parte no látex polimérico (MORRISON *et. al* 1992).

Immanuel, Wang e Bianco, (2008) apresentam a clara influência da distribuição do tamanho de partícula sobre a reologia do látex, assim como a sua relação indireta com o conteúdo de sólidos que também influencia a reologia do látex. No fluxograma apresentado na Figura (4), o autor procura demonstrar como é complexa a interação da distribuição do tamanho de partícula durante o mecanismo de polimerização.

Figura 4 Interações ocorridas durante o mecanismo de polimerização e as relações com a DTP



Fonte: Immanuel, Wang e Bianco (2008)

Em polimerização, a produção de polímeros com alto teor de sólidos é algo muitas vezes desejável, já que apresenta algumas vantagens, como o melhor aproveitamento do reator em termos de tempo e espaço, além de reduzir custos de armazenamento e transporte, podendo até melhorar algumas propriedades do produto para determinadas aplicações (MARINANGELO, 2010 apud BOUTTI, GRAILLAT e McKENNA, 2005).

Em uma distribuição monomodal de partículas grandes, à medida que se aumenta o teor de sólidos, ocorre uma redução significativa do espaço entre as partículas afetando a estabilidade da emulsão. Por outro lado, para o mesmo teor de sólidos, uma distribuição monomodal de

partículas pequenas apresenta menor distância entre as partículas, resultando em uma maior interação entre as partículas, apresentando por consequência maiores viscosidades. Nesse sentido, de acordo com Guyot *et. al* (2002) a produção de látex polimérico com alto conteúdo de sólidos requer um controle preciso da DTP.

Amaral *et. al* (2004) citam que produtos com maior teor de sólidos podem ser obtidos através de uma distribuição de tamanho de partícula larga ou bimodal. Em seu estudo, Guyot *et. al* (2002) obtiveram sucesso na produção de látex com teor de sólidos superior a 65%, sendo que os maiores teores de sólidos foram obtidos para distribuições bi e trimodais. Estes resultados ocorrem devido à possibilidade de partículas menores poderem se acomodar nos interstícios deixados por partículas maiores, havendo um maior empacotamento com relativo aumento da interação entre as partículas. Além disso, as partículas pequenas, preenchendo os interstícios entre as partículas grandes, facilitam o deslocamento dessas partículas maiores, agindo como uma espécie de lubrificante, reduzindo a viscosidade do látex (MARINANGELO, 2010).

Meyer e Keurentjes (2005) argumentam que a distribuição de tamanho de partícula, em muitas ocasiões, está intimamente relacionada ao aumento de escala e a viabilidade de operação de um reator de polimerização, visto que esta afeta fortemente a reologia do látex polimérico que, por sua vez influencia a taxa de remoção de calor, mistura, transferência de massa e estabilidade do látex.

O tipo do processo também tem um forte impacto no andamento da polimerização em emulsão. No caso dos processos em batelada ou batelada alimentada, um dos maiores problemas está associado ao fato de que produtos totalmente diferentes podem ser produzidos pela mesma formulação. Os reatores contínuos seriam aparentemente a solução para este problema de qualidade, porém apresentam desvantagens também, como a existência de um comportamento oscilatório devido à renucleação de partículas e à multiplicidade de estados estacionários para uma determinada faixa de condições de operação.

Segundo Prindle (1989), o comportamento oscilatório está associado às condições de operação, que fazem com que o processo de nucleação de novas partículas seja interrompido, e depois reiniciado. Assim, o desaparecimento das micelas pode ocorrer em função do crescimento das partículas poliméricas que, depois de retiradas do reator, provocam aumento da concentração de emulsificante livre e reaparecimento das micelas, proporcionando uma nova renucleação. Esse comportamento também pode ser observado durante a nucleação

homogênea, pois quanto menor a área superficial total das partículas, menor será a entrada de radicais oligoméricos nas partículas e, portanto, maior será a probabilidade deste radical atingir o tamanho crítico formando uma nova partícula, aumentando a nucleação homogênea. Entretanto, o novo aumento do número de partículas e da área superficial total leva a um decréscimo da nucleação homogênea, gerando um novo ciclo.

Um estudo dos efeitos de determinadas variáveis de operação sobre a dinâmica de um reator tubular contínuo com reciclo (reator *loop* contínuo) e mais especificamente sobre a DTP em um processo de polimerização heterogênea foi realizada por Araújo (1999) tendo verificado que o sistema apresenta multiplicidade de estados estacionários, fenômeno este atribuído ao processo de renucleação. Em relação às variáveis de operação o autor faz as seguintes considerações:

- ✓ Concentração de iniciador: látex estáveis de um polímero de alta massa molecular podem ser obtidos a altas taxas de polimerização se esta é iniciada usando uma alta concentração de iniciador, que é mais tarde reduzida para o valor desejado. A qualidade do produto é beneficiada pela redução da concentração de iniciador, já que a massa molecular do polímero aumenta e isto leva a um aumento da resistência ao atrito do filme formado a partir do látex. Além disso, reduções da concentração de iniciador diminuem a sensibilidade a água do filme formado.

- ✓ Tempo de residência: altas taxas de conversão de monômero em tempos de residência curtos, buscando atingir o máximo de produção, podem ser atingidas iniciando-se a reação com tempos de residência longos e reduzindo-os quando o estado pseudo-estacionário é atingido.

- ✓ Temperatura do reator: Quanto maior a temperatura, maior a taxa de formação de radicais e de polimerização, mas menor é o tempo que um determinado elemento de fluido reage, pois o iniciador é consumido rapidamente. Estes efeitos contrapostos, juntos com a baixa sensibilidade do sistema a altas conversões, talvez possam explicar porque a conversão não diminui ao reduzir a temperatura de reação.

- ✓ Concentração de emulsificante: Durante o início do processo, a superfície total das partículas atinge um máximo em torno de dois tempos de residência e depois decresce até o estado estacionário. Além disto, a viscosidade do látex também atinge o auge nesta faixa de tempo de residência. Isto sugere que látex com baixa concentração de emulsificante possa ser obtido iniciando-se o processo com uma

quantidade relativamente alta de emulsificante. Esta quantidade é reduzida, uma vez que o estado estacionário tenha sido atingido.

Em síntese, segundo Araújo (1999) é possível tirar proveito da dinâmica particular do reator *loop* e conduzir a reação de polimerização de tal forma que um estado pseudo-estacionário seja atingido, no qual altas conversões de monômero, em curtos tempos de residência, baixa temperatura e baixa concentração de iniciado, são obtidas. Basicamente, a estratégia é iniciar o processo com um valor relativamente alto de tempo de residência, temperatura e concentração de iniciador e mais tarde reduzi-los para os valores desejados. Vale à pena ressaltar que se estas condições finais fossem implementadas desde o início do processo seriam obtidas baixas conversões, e algumas vezes coagulação.

Além das considerações citadas, Capek e Potisk (1995) argumentam que o aumento da temperatura promove o aumento da taxa de formação de novas partículas, o que provoca a redução do tamanho médio e aumenta a concentração das partículas. Por fim, outros fatores como agitação, que afeta a taxa de coalescência, e a forma de alimentação do reator também exercem influência sobre a taxa de reação e consequentemente, sobre a DTP (ZUBITUR e ASUA, 2001).

Um dos desafios principais no monitoramento e obtenção de distribuição de tamanho de partícula nas condições desejáveis em reações de polimerização está na obtenção de medidas *on line*. Isto requer a aquisição de dados de distribuição de tamanho de partícula a uma frequência razoável para facilitar as ações corretivas.

Experimentalmente, estas distribuições podem ser determinadas por métodos que podem ser divididos em três classes:

✓ *Ensemble Methods* – destacando-se a *Dynamic Light Scattering* [DLS] e *Quasi-Elastic Light Scattering* [QELS] que mede movimentos *Brownianos* e os relaciona ao tamanho das partículas apresentando como principal desvantagem a baixa resolução e a dificuldade de determinar distribuições polidispersas.

✓ *Counting Methods* – com destaque para a *Transmission Electron Microscopy* [TEM] sendo este o mais preciso meio de determinar distribuições de tamanho de partícula em polimerização (GUGLIOTTA et. al 2009). Contudo, métodos microscópicos necessitam de uma preparação prévia da amostra e podem apresentar uma complexidade e tempo de análise excessivos.

✓ *Separation Methods* – dentre esse método, ressalta-se a Centrífuga Disco, sendo o método um dos mais usuais, mas para distribuições polidispersas e partículas muito pequenas (>100nm) o tempo de análise de cada amostra é relativamente alto (ARAÚJO, 1999),

e o *Capillary Hydrodynamic Fractionation* (CHFD) com desvantagens em relação à baixa resolução, a capacidade de realizar medidas somente em meio aquoso e a dificuldade em medir corretamente partículas que não sejam perfeitamente esféricas.

Apesar de avanços recentes, principalmente em técnicas de reflexão Ultravioleta, *Raman Spectroscopy* e *Near Infrared Spectroscopy* (REIS *et. al* 2002; CHICOMA, SAYER e GIUDICI, 2011; PINTO *et. al* 2011), e *Capillary Hydrodynamic Fractionation* (MARIZ, LEIZA e DE LA CAL, 2011; MALLIKARJUNAN, PUSHPAVANAM e IMMANUEL, 2010), dificuldades em determinar importantes propriedades relativas à qualidade e produtividade dos polímeros ainda persistem (MACHADO, LIMA e PINTO, 2007). Por conta disso, “sistemas inteligentes” têm sido desenvolvidos, baseados em modelos matemáticos e estatísticos em associação com técnicas computacionais para prever satisfatoriamente variáveis como a distribuição de tamanho de partícula durante o processo de polimerização.

2.3 MODELOS MATEMÁTICOS DA DISTRIBUIÇÃO DO TAMANHO DE PARTICULAS EM POLIMERIZAÇÃO

As tentativas iniciais para quantificar as variáveis da polimerização em emulsão eram modelos baseados em aproximações monodispersas, assumindo que todas as partículas tinham o mesmo volume. Apesar do sucesso em determinados casos, esse modelo apresenta certas limitações, como o fato de que cinética de polimerização é aproximada, visto que o número médio de radicais por partícula é, em geral, uma função não linear do volume das partículas e, conseqüentemente, a simplificação não é suficiente para descrever o sistema.

Atualmente, os modelos que descrevem a polimerização em emulsão que incluem equações do balanço populacional [EBP] não sofrem com este tipo de restrição sendo, portanto, mais adequados quando é importante a dinâmica de nucleação e crescimento e quando assumir simplificações não é aceitável, ou em casos onde é necessário considerar diferentes cinéticas, composições, etc. como função do tamanho de partícula (VALE e McKENNA, 2005).

Segundo Crowley *et. al* (2000), um modelo de balanço detalhado da população é uma ferramenta inestimável e necessária para a proposta de otimização que tem como objetivo identificar as condições operacionais ótimas que conduziriam a distribuição de tamanho de partícula do látex às condições desejadas.

Uma das primeiras discussões a respeito da modelagem de polimerização em emulsão foi proposta por Min e Ray (1974), que apresentam um modelo muito geral que inclui equações do balanço populacional para descrever distribuição de tamanho de partícula e distribuição de massa molecular em reatores de homo polimerização em emulsão. Morbidelli, Storti e Carra (1983) apresentaram um modelo matemático onde o mecanismo de nucleação pode ser micelar ou homogêneo. Alguns anos depois, Saldivar, Dafiniotis e Ray, (1998) estende este modelo para sistemas de copolimerização, incluindo um esquema cinético completo. Neste mesmo período, Coen *et. al* (1998) modelaram um sistema de polimerização de estireno em batelada em que são considerados os mecanismos de nucleação homogênea e micelar, balanço de radicais na fase aquosa e no interior das partículas e também a coalescência de partículas precursoras.

O balanço populacional para as partículas de polímero na polimerização em emulsão é determinado pela interação de três processos principais: a nucleação, o crescimento e a coalescência. A quebra das partículas, termo comumente encontrado em balanços populacionais, pode ser desconsiderada no caso da polimerização em emulsão, pois, devido ao tamanho reduzido e coesão das partículas, a sua quebra é incomum (MARINANGELO, 2010). O Apêndice A apresenta uma síntese da modelagem da distribuição do tamanho de partícula, fazendo uso das equações do balanço populacional.

Observando as equações que compõem o balanço populacional percebe-se que os mecanismos de nucleação das partículas, crescimento e coalescência são dependentes do sistema e requerem o conhecimento de vários parâmetros incertos que são difíceis de serem determinados. Segundo Zhang e Ray (2002), o processo de polimerização envolve muitos parâmetros que variam com o tempo e os “dados” do modelo (velocidade de reação, massa molecular, tamanho e número de partículas) sendo muito sensíveis a perturbações não mensuráveis (tipicamente resultantes da presença de traços de inibidores de polimerização ou outros componentes que podem estar presentes nos reagentes) e às condições iniciais, que nem sempre são conhecidas com precisão. Corroborando com esta afirmação Mariz, Leiza e De la Cal, (2011) atestam que a capacidade de predição dos modelos existentes ainda é limitada.

Genericamente, as equações do balanço populacional são descritas por um conjunto infinito de equações integro-diferenciais parciais hiperbólicas e admitir todas as situações envolvidas no processo de polimerização em emulsão tornam estas equações intratáveis. O

caminho para superar estas dificuldades é reduzir a dimensionalidade do problema fazendo algumas aproximações, como: desconsiderar a terminação das cadeias longas, partículas com dois ou mais radicais e efeitos da compartimentalização.

Geralmente a solução das equações do balanço populacional é obtida com a discretização da distribuição (função densidade populacional) em certo número de elementos discretos, resultando em um sistema não-linear composto de equações algébrico-diferenciais (ULIANA, 2007).

A solução numérica por discretização do balanço populacional envolve a aproximação da distribuição contínua por um conjunto de pontos ou classes discretas. A eficiência dos diferentes métodos de discretização na solução das equações do balanço populacional varia de problema para problema e depende dos fenômenos preponderantes (ALEXOPOULOS, ROUSSOS e KIPARISSIDES, 2004).

A desvantagem desses métodos é que, em geral, o número de pontos/classes N precisa ser grande resultando num problema de grande esforço computacional (ULIANA, 2007; IMMANUEL, WANG e BIANCO, 2008). Além disso, são relatadas algumas dificuldades na solução numérica do conjunto de equações integro-diferenciais parciais como a natureza hiperbólica destas equações levando a uma dispersão numérica (LE VEQUE, 2002); e ainda qualquer erro no cálculo de coeficientes como taxa de coalescência resultará em uma não-conservação da massa do polímero (RIGOPOULOS, 2003).

Discussões relevantes na formulação das equações do balanço populacional referentes à escolha das coordenadas internas; formulação das condições de contorno; aspectos da coagulação das partículas e suas aplicações podem ser encontrados em Ramkrishna (2000). Uma revisão bastante completa a respeito da resolução das equações do balanço populacional, através dos métodos numéricos, podem ser encontradas em Alexopoulos, Roussos e Kiparissides (2004) e Uliana (2007).

A partir do conhecimento mais detalhado da cinética de polimerização e da exploração de métodos numéricos na resolução dessas equações, muitos trabalhos têm sido propostos para a modelagem e controle da distribuição do tamanho de partícula em reatores de polimerização em emulsão. A Tabela 2 apresenta alguns dos modelos baseados em equações de balanço populacional mais significativos, trazendo a contribuição dos mesmos em relação à determinação das DTPs.

Tabela 2 Modelos de balanço populacional em polimerização em emulsão

AUTOR(ES)	CONTRIBUIÇÃO
Araújo <i>et. al</i> (2001a)	Descrever a evolução das DTPs durante a copolimerização em emulsão de VAc e Veova 10 em reator <i>loop</i> contínuo, com comparação com dados experimentais. A técnica de colocação ortogonal em elementos finitos é utilizada na resolução do modelo. Segundo os autores, a DTP calculada conseguiu acompanhar os golpes de nucleação com distribuições bimodais.
Kiparissides, Achilias e Frantzikinakis (2002)	Quantificar o efeito da concentração de oxigênio sobre a taxa de polimerização e DTP em um reação de polimerização em emulsão de cloreto de vinil em batelada. Diferenças finitas atrasadas são usadas para resolver o modelo, havendo nesse caso, comparação dos resultados do modelo com dados experimentais.
Zeaiter <i>et. al</i> (2002).	Promover o cálculo da DTP e da distribuição de peso molecular, com validação do modelo com dados experimental. Reator semi-batelada.
Sood, (2004)	Investigar o comportamento da DTP em reatores de polimerização em batelada e semi-batelada. Colocação ortogonal em elementos finitos é usada para resolver o modelo, sendo este usado para investigar os efeitos do controle operacional da quantidade inicial de iniciador e surfactante sobre a DTP. O fenômeno da coalescência é ignorado na formulação do modelo.
Vale e McKenna, (2007)	Propõem uma solução numérica para as EBP. A técnica baseia-se no método de volumes finitos e utiliza esquemas de alta-resolução e uma técnica de pivô fixo generalizada para diferenciar as condições de crescimento e de agregação, respectivamente. No entanto, não realiza a validação da técnica com dados experimentais.
Sweetman <i>et. al</i> (2008)	Determinar DTP e distribuição de peso molecular em um sistema reacional de copolimerização de vinil acetato e butil acrilato, com validação experimental. O balanço populacional de DTP é resolvido usando um algoritmo de decomposição, que resolve separadamente a taxa de nucleação, crescimento e coagulação das partículas, usando-os para atualizar a DTP em cada intervalo de tempo.

Mallikarjunan <i>et. al</i> (2010)	Realizar a predição da DTP em copolimerização em acetato de vinila e butil acrilato, conduzido em reator tanque agitado semi-batelado, objetivando corrigir o atraso da renucleação e corrigir a taxa de crescimento das partículas. No entanto, sem discutir aspectos relativos à conservação de massa, importantes quando há alterações na taxa de crescimento de partículas.
Hosseini, Bouaswaig e Engell (2012)	Propõem a inclusão de um termo estocástico no modelo de crescimento das partículas, com a intenção de representar a heterogeneidade da evolução das partículas. Dados experimentais da polimerização de estireno em regime semi-batelada são utilizados na validação do modelo proposto.

Fonte: Desenvolvido pelo autor

Autores como Vale e McKenna, (2005) argumentam que o desenvolvimento de modelos de balanço populacional para descrever o comportamento da DTP em polimerização em emulsão é considerado ainda um desafio. Zeaiter, Romagnoli e Gomes, (2006) apresentam afirmação semelhante quanto ao controle de DTP durante o processo de polimerização em emulsão, segundo eles dificuldades associadas à determinação em tempo real da DTP, em conjunto com os mecanismos complexos envolvidos em sistemas de polimerização em emulsão, em geral, limitam as opções de operação e fazem do problema do controle uma tarefa complexa. Em tais casos, as estratégias de controle convencionais não conseguem garantir uma qualidade consistente do produto.

Immanuel *et. al* (2008) avaliam que modelos que incluem equações de balanço populacional, para a determinação da DTP requerem um esforço computacional absolutamente inadequado as necessidades de uma estratégia de monitoramento e controle das condições desejadas do látex polimérico *on line*. Mais recentemente, Hosseini, Bouaswaig e Engell (2012) argumentam que os modelos de balanço populacional clássicos são incapazes de prever a evolução e amplitude das DTPs experimentais quando um método de discretização de alta resolução é usado para suprir os erros numéricos.

Nos últimos tempos, no entanto, os avanços no processo de entendimento da reação, modelagem matemática, tecnologia em soft-sensor e modelagem baseada em técnicas de controle tem propiciado possibilidades de promover grandes melhorias nas condições de

operação do processo e na qualidade do produto. A abordagem viável parece combinar essas ferramentas poderosas em uma única estratégia. (ZEAITER, ROMAGNOLI e GOMES, 2006).

Por conta disso, a utilização de ferramentas computacionais como as Redes Neurais Artificiais na inferência das DTPs, pode representar um avanço importante na minimização do problema da determinação da DTP durante a reação de polimerização em emulsão

2.4 REDES NEURAIIS ARTIFICIAIS

2.4.1 Princípios

RNs Artificiais são modelos matemáticos que se assemelham às estruturas neurais biológicas e que tem capacidade computacional adquirida por meio de aprendizado e generalização (BRAGA *et. al* 2000).

De acordo com Haykin (2001):

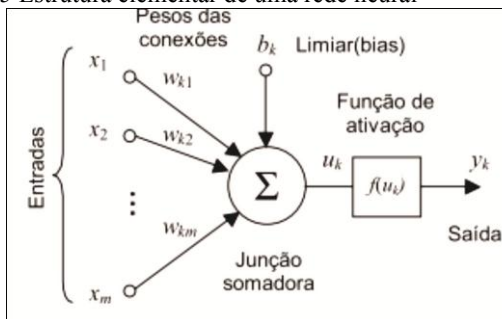
Uma rede neural é um processador, maciça e paralelamente distribuído, constituído de unidades de processamento simples, que tem a propensão natural para armazenar conhecimento experimental e torná-lo disponível para uso. Ela se assemelha ao cérebro em dois aspectos: i) o conhecimento é adquirido pela rede a partir de seu ambiente através de um processo de aprendizagem; ii) forças de conexão entre neurônios, conhecidas como pesos sinápticos, são utilizadas para armazenar o conhecimento adquirido.

As RNs Artificiais congregam seu conhecimento através da detecção de padrões e das relações entre informações dos dados com os quais elas são treinadas. Portanto os resultados de saída de uma rede neural são obtidos através da experiência e não via programação direta (GIORDANI, 2003).

A habilidade das RNs vem da conexão entre os neurônios na forma de uma rede. Cada elemento processador possui uma entrada que por sua vez, possui um peso, uma função de transferência e uma saída, como pode ser visto na Figura (5). O comportamento da rede é determinado pelas funções de transferência entre os neurônios, através da regra de aprendizagem e pela sua arquitetura. Os pesos dos neurônios são os parâmetros ajustáveis e, nesse sentido, uma rede pode ser considerada um sistema parametrizado.

Uma grande vantagem das RNs é o fato delas serem capazes de usar algumas informações que, a priori, estão ocultas nos dados (mas elas não são capazes de extraí-las). O processo de capturar essas informações desconhecidas é chamado aprendizado ou treinamento da rede neural. Em matemática formal, o aprender da rede significa ajustar os pesos sinápticos dos neurônios de tal maneira que determinadas condições sejam satisfeitas (HAYKIN, 2001).

Figura 5 Estrutura elementar de uma rede neural



Fonte: Haykin (2001)

As RN têm como principal característica a capacidade de adaptação a comportamentos não-lineares, sendo capazes de:

- ✓ Aprender a partir da interação com o meio ambiente, sem restrições para capturar qualquer tipo de relação entre os padrões de informação mesmo se muitas informações forem administradas;

- ✓ Generalizar as informações aprendidas para situações similares nunca vistas antes, e

- ✓ Possuir um bom grau de tolerância ao erro, principalmente devido à sua intrínseca arquitetura em paralelo. Estas propriedades tornam as RNs um recurso interessante na solução de problemas da Engenharia Química, onde a não-linearidade é quase uma constante (RALLO *et. al* 2002).

Contudo, de acordo com Haykin (2001), na tentativa de generalizar, uma RN pode se especializar demasiadamente em relação aos exemplos contidos na base de aprendizado. Este tipo de comportamento gera um problema de aprendizado conhecido como superaprendizado (*over-training/over-fitting*) que pode ocorrer quando na arquitetura da RN o número de camadas e/ou neurônios for muito grande. Nessa situação, pode-se dizer também que está ocorrendo uma sobre parametrização da RN, pois o número de parâmetros ou pesos sinápticos fica elevado devido à grande arquitetura assumida. Outra

situação é a rede não conseguir generalizar, gerando o problema chamado de subajuste (*under-fitting*), o qual pode ocorrer quando o número de camadas e/ou neurônios for muito pequeno. A determinação do número de ciclos de treinamento, de acordo com um critério de parada específico, também dificulta o treinamento de RNs (BASHEER e HAJMEER, 2000).

2.4.2 Redes Neurais Multi-Layer Perceptron

O *perceptron* de múltiplas camadas (*Multi-Layer Perceptron* - MLP) é uma rede do tipo *perceptron* com, pelo menos, uma camada intermediária. Trata-se de uma generalização do conceito publicado no livro *Perceptrons* (MINSKY e PAPERT, 1969), onde foi proposta a rede *perceptron* com múltiplas camadas como uma tentativa de superar as limitações encontradas no *perceptron* simples. Nesse sentido, tornou-se necessário desenvolver algoritmos de treinamento mais sofisticados capazes de definir de forma automática os pesos para este tipo de rede (VON ZUBEN e CASTRO, 2010).

De acordo com Haykin (2001) uma rede neural MLP típica possui três características distintas:

- ✓ Os neurônios das camadas intermediárias e os da camada de saída possuem uma função de ativação não-linear do tipo sigmoidal, função logística ou tangente hiperbólica;

- ✓ A rede possui uma ou mais camadas de neurônios ocultos, cada camada adiciona um poder maior em relação às redes perceptron de camada única, que classifica apenas padrões linearmente separáveis, sendo os neurônios ocultos responsáveis por capturar a não-linearidade dos dados;

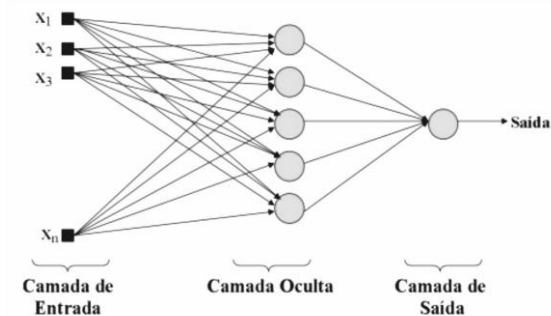
- ✓ A rede possui um alto grau de conectividade, determinado pelas sinapses da rede.

Na implementação de uma rede neural MLP é importante perceber que o número de neurônios em cada camada, a quantidade de camadas intermediárias na rede, o número de conexões, a função de ativação usada e a inicialização dos pesos são fatores importantes para a definição da arquitetura da rede e terão influência no processo de treinamento da rede neural. Segundo Basheer e Hajmeer (2000) a metodologia de implementação de rede neural mais usual é aquela onde estes fatores são determinados por tentativa e erro, ou seja, são atribuídos vários valores distintos aos parâmetros e, analisando os resultados obtidos, a melhor configuração é escolhida. Dentre esses parâmetros, o número de camadas ocultas e o número de neurônios nas

camadas ocultas possuem um maior grau de dificuldade para o ajuste dos valores.

A Figura (6) mostra a arquitetura de uma rede neural MLP com uma camada oculta, sendo esta uma rede progressiva [*feedforward*] visto que as saídas dos neurônios em qualquer camada particular se conectam unicamente às entradas dos neurônios da camada seguinte, sem a presença de laços de realimentação.

Figura 6 Estrutura básica de uma Rede Neural Multi-layer Perceptron



Fonte: Haykin (2001)

Na forma simples de uma rede neural artificial, cada neurônio é conectado a outros neurônios de uma camada anterior através de pesos. O neurônio recebe a informação dos outros neurônios da camada anterior através de suas conexões de entrada. Primeiramente, estas informações são somadas de maneira ponderada pelos pesos associados às conexões e, então, o resultado é o argumento de uma função de ativação, sendo assim passado para o próximo neurônio. Além disso, cada neurônio (com exceção das unidades da camada de entrada) possui um “bias” que pode ser interpretado como um peso aplicado a uma pseudo entrada de valor igual a 1 (KALOGIROU, 2003), ou seja, representa o sinal que mantém o neurônio ativo quando as variáveis de entradas são nulas.

De acordo com Haykin (2001), as RNs artificiais podem ser representadas pelas Equações (1) e (2):

$$y_k = f\left(\sum_i^n x_i \cdot w_{ik}^h + b_k\right) \quad (1)$$

$$Z_p = f \left(\sum_j^m y_i \cdot w_{kp}^o + b_p \right) \quad (2)$$

onde:

y_k : saída do k-ésimo neurônios da camada interna;

x_i : i-ésima entrada da rede neural;

Z_p : saída do p-ésimo neurônio da camada de saída da rede neural

b_k e b_p : —ais” do k-ésimo neurônio da camada interna e do p-ésimo neurônio da camada de saída;

w_{ik}^h : peso da conexão do i-ésimo neurônio da camada de entrada para o k-ésimo neurônio da camada interna

w_{kp}^o : peso da conexão do k-ésimo neurônio da camada interna para o p-ésimo neurônio da camada de saída;

f : função de ativação

n : número de entradas

m : número de unidades internas

O processo de treinamento supervisionado de RNs MLP é equivalente a um problema de otimização não-linear irrestrito, definidos em um espaço S de busca da solução definido em função de um conjunto de variáveis de decisão X_i , com $i = 1, 2, 3, \dots, n$. Estas variáveis apresentam domínio discreto e/ou contínuo possuindo uma série de restrições. Por fim, os problemas possuem uma função objetiva, também denominada função erro, $f: S \rightarrow R^+$, a qual é atribuída um custo para cada solução em S , sendo que o objetivo é encontrar uma solução $s \in S$ tal que $f(s) \leq f(s')$, para todo e qualquer $s' \in S$ no caso da minimização de funções objetivo (BLUM e MERKLE, 2008).

Devido à sua facilidade de implementação e o comprovado histórico de sucessos, o algoritmo *backpropagation*, baseado numa variação da técnica de gradiente descendente, foi inicialmente utilizado na maioria dos trabalhos que tratam de treinamento de RNs (RUMELHART e MCCLELLAND, 1986; SEXTON, DORSEY e JOHNSON, 1998).

No entanto, a característica determinística do *backpropagation* se reflete em alguns problemas freqüentemente encontrados na utilização desse tipo de algoritmo. O algoritmo *backpropagation* pode facilmente ficar preso em mínimos locais, especialmente para problemas que possuem comportamento altamente não linear sendo, nesse caso, muito dependente das escolhas dos valores iniciais dos pesos de conexão de rede (GORI e TESI, 1992). Além disso, a velocidade de convergência do algoritmo *backpropagation* é muito lenta, sendo esta dependente dos

parâmetros do algoritmo, tais como a taxa de aprendizagem e do momento (HAYKIN, 2001).

Para melhorar o desempenho do algoritmo original, esforços foram realizados explorando os seguintes fatores: seleção da melhor função de ativação (VAN OOOYEN e NIENHUIS, 1992); seleção de taxa de aprendizagem dinâmica e momento (JACOBS, 1988; WEIRS, 1991). No entanto, essas melhorias não removeram as desvantagens do algoritmo *backpropagation* (SEXTON, DORSEY e JOHNSON, 1998; WANG, 1995; ZHANG *et. al* 2007).

Sendo uma variação do Método de *Newton*, o algoritmo de *Levenberg Marquart*, também apresenta bom desempenho na minimização de funções não lineares (HAGAN, DEMUTH e BEALE, 2003). O algoritmo de *Levenberg Marquart* foi aplicado no treinamento de RNs por autores como Singh *et. al* (2007), Zhao e Xu (2004), Wilamowski *et. al* (2001). No entanto, modelos de RNs são tipicamente não lineares possuindo, portanto, muitos mínimos locais, o que não favorece os métodos de otimização baseados em gradiente que ficam refém de uma inicialização fortuita (HERRERA e ZHANG, 2009).

Uma alternativa a esse problema é a utilização de algoritmos estocásticos. Esses algoritmos, como o *Particle Swarm Optimization* [PSO] e o *Genetic Algorithm* [GA] podem ser caracterizados como uma técnica heurística de otimização global (LINDEN, 2006). São usualmente métodos de aproximação, onde a garantia de apresentar a solução global é sacrificada por uma solução local confiável em um tempo computacional significativamente reduzido (BLUM e MERKLE, 2008), sendo caracterizados pela realização de um grande número de avaliações da função objetivo em toda a região de busca, de forma a aumentar a probabilidade de que a solução local encontrada represente o mínimo global da função objetivo. Vale ressaltar também que o caráter aleatório do procedimento de busca, não dependendo de uma estimativa inicial, em muitos casos, evita que o mesmo fique retido em um mínimo local. Além disso, esses métodos não usam a derivada das funções, o que no caso de funções complexas é uma grande vantagem (SCHWAAB, 2005).

O GA é um algoritmo iterativo, baseados em inteligência computacional, sendo o mais utilizado na otimização de RNs (ZANCHETTIN e LUDERMIR, 2007). Autores como Stepniewski e Keane (1997) e Yao (1999) fizeram uso dos algoritmos genéticos na otimização de arquiteturas e pesos sinápticos de RNs MLP. O algoritmo genético foi utilizado com sucesso por Sexton, Dorsey e Johnson, (1998) no treinamento de RNs, em comparação com o *backpropagation* e o

Simulated Annealing. Este mesmo algoritmo foi utilizado combinado ao Método de *Newton* Modificado em um algoritmo de dois estágios no treinamento de RNs *feedforward* com eficiência (SU, YANG e HUANG, 2011).

Contudo, autores como Kiranyaz *et. al* (2009); Jiao, Lian e Gu (2008), afirmam que a implementação do GA é mais complexa que o PSO, visto que este possui operadores evolutivos mais sensíveis como a mutação, seleção e cruzamento.

Assim, o *Particle Swarm Optimization* tem sido aplicado com sucesso no treinamento de RNs *feed-forward* (MEISSNER *et. al* 2006; YU *et. al* 2008; VAN DEN BERGH, 1999) tendo se mostrado superior ao tradicional *Backpropagation* em termos de precisão e rapidez. Carvalho e Ludermir (2007), aplicaram o PSO na otimização de arquiteturas e pesos sinápticos de RNs MLP, sendo que esta metodologia foi apresentada inicialmente por Zhang, Shao e Li, (2000). Outro trabalho relevante é apresentado por KIRANYAZ, *et. al* (2009) que propõem uma rede neural artificial evolucionária, onde o projeto da rede é definida de forma automática pelo algoritmo PSO.

Ahmadi (2012) combina o PSO com o algoritmo *Backpropagation* para formar um algoritmo híbrido no treinamento de RNs artificiais. Outros algoritmos híbridos com o mesmo objetivo são propostos por Wang e Lu (2006) e Leung, Tang e Wong (2012). Dessa forma, a utilização do algoritmo PSO vem se consolidando como uma das estratégias mais eficientes na otimização de arquiteturas e pesos sinápticos de RNs artificiais.

Por fim, algoritmos de busca global, baseados em inteligência computacional, tem sistematicamente se mostrado eficientes na otimização de RNs artificiais, não havendo uma convergência de opiniões quanto à escolha do mais eficiente e preciso. De forma geral, certas configurações de algoritmos se adequam melhor a situações específicas, seja na realização de treinamento de tipos de RNs diferentes e ou na finalidade a qual está destinada esta rede neural, não havendo prevalência em termos de desempenho de nenhum deles.

2.4.3 Aplicações de Redes Neurais ao Processo de Polimerização em Emulsão

A estimação de parâmetros cinéticos da reação de polimerização envolve o desenvolvimento de uma ferramenta matemática derivada de dados experimentais e, uma vez estabelecida pode, de forma significativa, reduzir o trabalho experimental envolvido na concepção de novos compostos poliméricos. As RNs podem ser utilizadas para

resolver uma ampla variedade de problemas complexos, em função da sua natureza, onde o aprendizado ocorre através de exemplos. Por isso, podem ser treinadas para encontrar soluções de problemas não lineares, multidimensionais e que possuem relações funcionais sem suposição prévia sobre a sua natureza (ZHANG e FRIEDRICH, 2003).

Outras aplicações são pertinentes, como quando a predição da qualidade do produto final é uma tarefa que se torna difícil, pois não é possível a análise *on line* das propriedades que compõem o produto da reação de polimerização, prejudicando qualquer estratégia de monitoramento e controle desse processo. Uma alternativa nestas situações é a implementação de estimadores *on-line*, baseados em RNs, para inferir a qualidade desses produtos através de informações do processo (MARTIN, 1997). Na última década, os trabalhos que se referem à otimização e controle de reatores de polimerização, utilizando-se das vantagens oferecidas pelas RNs, têm se multiplicado (EPKO e MUJTABA, 2008; GONZAGA *et. al* 2009; CAO *et. al* 2007; CONSTANT, 2007; HERRERA e ZHANG, 2009; GAO *et. al* 2013)

Em relação às variáveis do processo, a temperatura da reação é citada constantemente pelos autores como um fator crítico na estratégia de controle da polimerização em emulsão, visto que a taxa de geração de calor (reação exotérmica) não pode exceder a taxa de retirada de calor. Esta alteração no perfil de temperatura da reação pode interferir na qualidade do produto requerido e representar riscos quanto à segurança do processo (ETCHELLS, SNEE e WILDAY, 2000).

Observando esta limitação, Kuroda e Kim (2002) realizaram um trabalho no qual obtiveram os perfis de temperatura de polimerizações em emulsão em batelada por meio de um modelo de RNs, utilizando dados de temperatura do banho termostático que alimenta a camisa do reator e de sua respectiva vazão, obtendo sucesso especialmente no caso de controle preditivo. Contudo, os autores afirmam que o número de variáveis de entrada teve um efeito significativo na eficiência de aprendizagem da rede e na capacidade de adaptação.

Cao *et. al* (2007) discutem a implementação de um modelo da distribuição de temperatura em um reator tubular de polimerização desenvolvido através de RNs *B-spline*, relacionando as variáveis do processo a temperatura em cada ponto do reator no intuito de monitorar a distribuição da massa molecular do polímero através da rede neural. Ekpo e Mujtaba apresentaram em 2008 um trabalho sobre controle ótimo da reação de polimerização em batelada de metacrilato de metila, onde o processo de polimerização é descrito através do modelo cinético e das equações de balanço, sendo que a rede neural é utilizada para

estimar o calor liberado na reação no intuito de preservar características moleculares do polímero.

Tsen *et. al* (1996) desenvolveram uma metodologia baseada em modelos híbridos usando RNs e aplicaram à polimerização em emulsão de acetato de vinila. De acordo com os autores, foi possível controlar a distribuição de tamanho de partículas poliméricas, massa molecular médio e tempo em um reator batelada, demonstrando uma melhor capacidade de predição para a massa molecular em relação ao emprego de um modelo teórico. De acordo com Assis (2001) há boas vantagens no emprego de modelos híbridos, pois eles possuem certo conhecimento físico do processo, mas vários parâmetros carecem de determinação a partir de dados observados sendo que nesses casos as RNs podem ser utilizadas na descrição das partes desconhecidas do processo.

O trabalho de Gonzaga *et. al* (2009) apresentou uma rede neural *feed-forward* para desenvolver um estimador *on line* da viscosidade em uma estratégia de controle em um processo industrial de polimerização, que se mostrou efetiva. Nessa mesma linha, Tian, Zhang e Morris (2002) aplicaram combinações recorrentes de RNs para caracterizar o efeito gel, que representa uma das partes mais difíceis na modelagem da polimerização via radicais livres. A viscosidade, assim como o teor de sólidos do látex e a poli-dispersividade são propriedades modeladas no trabalho de Delfa, Olivieri e Boschetti (2009) com o auxílio de RNs. A viscosidade também pode ser usada como medida para inferir o valor da massa molecular do polímero, mas a precisão das medidas pode ser relativamente ruim (DE LANEY e OLIVER, 2002).

Uma aplicação das RNs artificiais na determinação das condições ótimas de reação no desenvolvimento de resinas de polietileno foi realizada por Fernandes e Lona (2002), utilizado um reator de leito fluidizado. Fernandes *et. al* (2004) utilizaram a mesma estratégia na determinação das condições ótimas de operação na polimerização em emulsão de acetato de vinila. Com objetivos semelhantes, Hanai *et. al* (2003) utilizaram uma rede neural artificial do tipo *Fuzzy Neural Network* combinada ao algoritmo genético para determinar as condições iniciais ideais de reação de preparação do poli-butadieno com características físico-químicas pré-determinadas.

A proposta de Giordani (2003) utiliza RNs artificiais associadas à calorimetria, como técnica de inferência de conversões individuais de copolimerizações em emulsão, em substituição ao observador de estado. Em um trabalho mais amplo, Araújo *et. al* (2001b) realizam a inferência de conversão e diâmetro médio de partícula a partir de várias variáveis importantes do processo em uma polimerização em reator loop contínuo,

apresentando boa correlação entre dados experimentais e calculados e demonstrando a eficiência das RNs.

No intuito de tornar viável o controle do processo de polimerização em emulsão, Zhang *et. al* (1998) desenvolveram uma técnica robusta para estimar propriedades dos polímeros utilizando representações de RNs combinadas com o estimador *on line* com o intuito de prever propriedades do poli-metacrilato de metila em um reator de polimerização em batelada. Os mesmos autores utilizaram as RNs combinadas na predição da quantidade efetiva de iniciador para o processo e do coeficiente de transferência de calor efetivo da reação, sendo essas variáveis utilizadas para determinação da quantidade de impurezas e resíduos resultantes do processo de polimerização (ZHANG *et. al*1999). Em 2008, Zhang apresentou em seu artigo uma estratégia de controle ótimo de reatores de polimerização em batelada, onde abdica de modelos fenomenológicos e procura simular o processo através de várias RNs combinadas. Estas RNs foram utilizadas para prever variáveis essenciais para a estratégia de controle como a massa molecular média, a conversão e a temperatura do reator. As RNs combinadas também foram utilizadas por Tian, Zhang e Morris (2002) sendo sua eficiência comparada com RNs MLP, mostrando a superioridade da performance das RNs combinadas na inferência de variáveis do processo de polimerização.

Uma síntese de alguns procedimentos bem difundidos de inferência através de RNs associados a alguns casos complexos na área de polimerização são discutidos por Fernandes e Lona (2005).

No entanto, não há registros de trabalhos que utilizam as RNs na estimação da DTP durante uma reação de polimerização em meio disperso. O trabalho que mais se aproxima dessa proposta é o de Gugliotta *et. al* (2009) que fazem uso da função lognormal unimodal para modelar distribuição de tamanho de partículas de látex polimérico com características assimétricas em um processo simulado de polimerização de poliestireno. O objetivo dos autores é estimar DTP através do reconhecimento de padrões por meio de uma rede neural alimentada com uma combinação de medidas de DLS (*Dynamic Light Scattering*) e ELS (*Elastic Light Scattering*).

2.5 PARTICLE SWARM OPTIMIZATION

O *Particle Swarm Optimization* (Otimização por Enxame de Partículas) é uma técnica de computação evolucionária (KENNEDY e EBERHART, 1995) inspirada no comportamento social de espécies biológicas de enxames (aves, peixes e ou formigas). A teoria

fundamental do *swarm intelligence* sugere que mentes e culturas são afetadas por suas interações sociais, de forma que sistemas compostos por agentes pouco inteligentes e com capacidade individual limitada, possam ser capazes de apresentar comportamentos coletivos inteligentes ao interagirem entre si e com o meio ambiente (WHITE e PAGUREK, 1998).

Segundo Jiao, Lian e Gu (2008), uma partícula i em uma iteração k se desloca através do espaço solução com dois atributos: o vetor posição atual dentro de um espaço de busca N -dimensional $X_i^k = (x_1^k, x_2^k, \dots, x_N^k)$ do problema, com $x_n^{\min} \leq x_n^k \leq x_n^{\max}$ para todo N , onde x_n^{\max} e x_n^{\min} são os limites da coordenada n ; e o vetor velocidade, representada por $V_i^k = (v_1^k, v_2^k, \dots, v_N^k)$ nesse mesmo espaço N -Dimensional do problema.

O vetor posição e o vetor velocidade são inicializados randomicamente de acordo com as Equações (3) e (4), onde os números randômicos (r_n) gerados obedecem a uma distribuição uniforme no intervalo $(-1, 1)$.

$$X_i^k = (X_n^{\max} - X_n^{\min})r_n + X_n^{\min} \quad (3)$$

$$V_i^k = (2r_n - 1)V_n^{\max} \quad (4)$$

Kennedy (1997) propõe o controle da velocidade máxima com o propósito de evitar a divergência do vetor posição por excesso de velocidade, usualmente definido pela Equação (5):

$$V_{\max} = (X_n^{\max} - X_n^{\min})/2 \quad (5)$$

A cada iteração a velocidade e posição de todas as partículas são atualizadas de acordo com as Equações (6) e (7):

$$X_i^{k+1} = X_i^k + V_i^k \quad (6)$$

$$V_i^{k+1} = wV_i^k + c_1r_n(P_i^k - X_i^k) + c_2r_n(P_g^k - X_i^k) \quad (7)$$

onde P_i^k é a melhor posição encontrada pela partícula i até a k -ésima iteração, sendo definido por:

$$P_i^k = \begin{cases} X_i^k, & \text{se } f(X_i^k) < f(P_i^{k-1}) \\ X_i^{k-1}, & \text{se } f(X_i^k) \geq f(P_i^{k-1}) \end{cases} \quad (8)$$

enquanto P_g^k é a melhor posição encontrada por todo o grupo até a k -ésima iteração, sendo definida por:

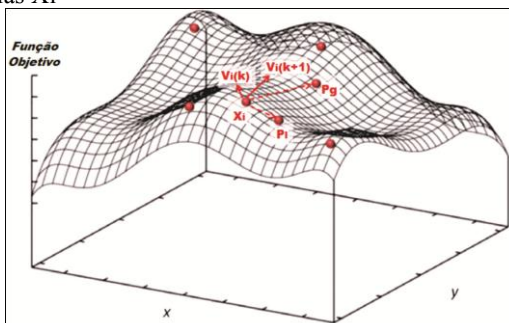
$$P_g^k = \arg \min \{f(P_1^k), f(P_2^k), \dots, f(P_n^k)\}$$

onde n é o número total de partículas.

Os resultados são atualizados segundo o seguinte critério: Se $f(X_i^k) < f(P_g^k)$ então, $f(P_g^k) = f(X_i^k)$.

A Equação (6) realiza a atualização da posição da partícula na iteração k e a Equação (7) define o vetor velocidade da partícula na iteração k , sendo afetada diretamente por dois vetores: o primeiro é a melhor posição encontrada pela partícula até o momento, indicada por P_i , e o segundo indicado por P_g , que é a melhor posição encontrada até o momento por qualquer indivíduo da população. A Figura (7) traz uma representação dos vetores que exercem influência sobre movimento realizado pelo vetor posição a cada iteração.

Figura 7 Representação vetorial das influências sobre o movimento das partículas X_i



Fonte: Adaptado de Chen, Chang e Ho (2011)

De acordo com Blum e Merkle, (2008) os termos que compõem o vetor velocidade tem a seguinte representação: o termo wV_i denominado momento, representa a tendência da partícula em continuar seu sentido atual. Já o termo $c_1 r_n(p_i^k - x_i^k)$, representa a distância entre a partícula i e sua melhor posição até a k -ésima interação. Este termo é chamado de componente cognitivo, pois ressalta a tendência da partícula em retornar a sua própria melhor posição encontrada até o momento, assumindo um papel de promover uma dispersão da busca em várias regiões do espaço

solução. De forma similar, o termo $c_2 r_n(p_g^k - x_i^k)$ representa a distância entre a partícula i e a melhor posição encontrada pela população até a k -ésima iteração. Este termo é chamado de componente social, pois ressalta a atração da partícula em relação à melhor posição encontrada até agora por todas as partículas, contribuindo para a convergência da partícula.

De forma específica, o parâmetro w conhecido como peso inercial, foi introduzido por Shi e Eberhart (1998) e hoje é comumente definido por uma função dinâmica. Este parâmetro tem como finalidade melhorar as habilidades de exploração do algoritmo, alterando o termo de momento. Inicialmente mais elevado, o parâmetro favorece a exploração do espaço solução. No decorrer das iterações, a redução do parâmetro favorece a convergência do algoritmo para o mínimo global, aumentando sua capacidade exploratória. Sua implementação reduziu a importância do controle da velocidade máxima (POLI, KENNEDY e BLACKWELL, 2007). Os parâmetros c_1 e c_2 são coeficientes de aceleração que ponderam a tendência do vetor posição mover-se na direção de P_i e P_g .

Desde o surgimento da técnica proposta por KENNEDY e EBERHART em 1995 o algoritmo PSO passou por uma série de adequações que vem conferindo mais eficiência e confiabilidade ao algoritmo. Diversos trabalhos se encontram disponíveis na literatura, propondo modificações em diferentes características do algoritmo, sendo os mais relevantes apresentados na Tabela (3).

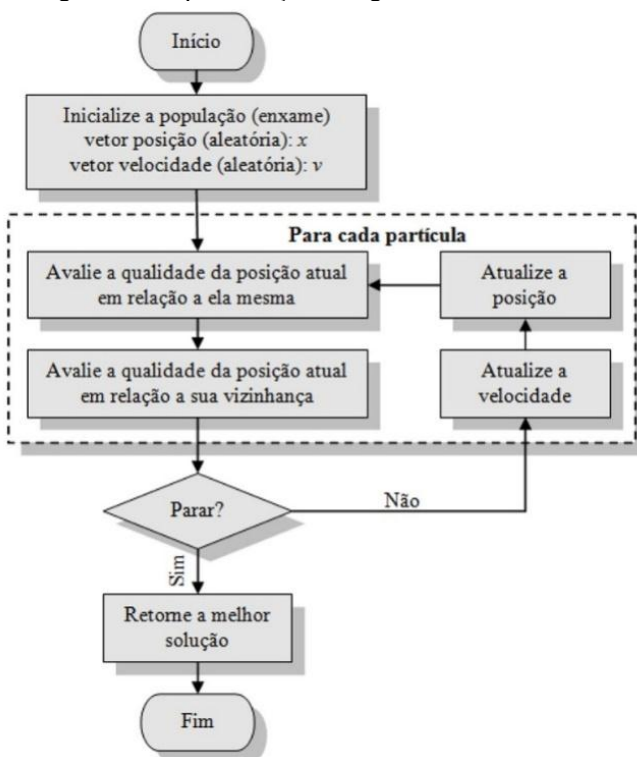
Tabela 3 Trabalhos que contribuíram para a evolução do PSO

Característica	Principais Trabalhos
Peso inercial	Shi e Eberhart (1998); Shi e Eberhart (1999); Jiao, Lian e Gu (2008); Liu <i>et. al</i> (2005), Chatterjee e Siarry (2004), Zhan <i>et. al</i> (2009)
Inicialização	Richards e Ventura (2003); Pant, Radha e Singh (2007).
Topologia	Kennedy e Mendes (2002); Handan, (2008); Katare, Kalos e West (2004); Omran, (2005)
Parâmetros	Trelea (2003); Van den Bergh e Engelbrecht (2006); Parsopoulos e Vrahatis (2006).
Diversidade	Olorunda e Engelbrecht (2008); Zhan <i>et. al</i> (2009); Lenin e Mohan, (2006); Chen, Chang e Ho (2011); Radha e Singh, (2007)
Híbridos	Thomas e Reed (2009); Pan e Zheng (2008); Katare, Kalos e West (2004); Thangaraj <i>et. al</i> (2011)

Fonte: Desenvolvido pelo autor

A Figura (8) apresenta um fluxograma da implementação do algoritmo PSO.

Figura 8 Fluxograma de implementação do algoritmo PSO



Fonte: Desenvolvido pelo autor

2.6 DISTRIBUIÇÕES DE PROBABILIDADE

2.6.1 Princípios

Uma distribuição de probabilidade ou distribuição de frequências relativas é uma síntese do agrupamento de dados em um espaço amostral e mostra a proporção das vezes em que a variável aleatória tende a assumir cada um dos diversos valores.

Em estatística, a função densidade de probabilidade $f(x)$ é utilizada para representar a distribuição de probabilidade caso a variável aleatória seja contínua, tendo ela limites de uma infinita sequência de

distribuições discretas, cujas variáveis formam um conjunto não enumerável de valores (ANDERSON; SWEENEY; WILLIAMS, 2008).

Tais distribuições devem atender às seguintes condições:

- $f(x) \geq 0$, para todos os valores de X , pois não existe probabilidade negativa;

- $\int_{-\infty}^{+\infty} f(x)dx = 1$, o que determina que a área entre a curva

representativa da função $f(x)$ em todo intervalo seja igual a 1

- $\int_a^b f(x)dx = P(a < X \leq b)$, $b > a$, ou seja, a probabilidade da

variável aleatória X assumir um valor em um intervalo será dada pela integral da função nesse intervalo. Geometricamente fica estabelecido que a probabilidade em um determinado intervalo corresponda à área determinada sob a região plana delimitada pela função nesse intervalo.

Uma propriedade decorrente da definição de função densidade de probabilidade é de que se uma função $f(x)$ satisfizer as condições

$f(x) \geq 0$ para todo x , $\int_{-\infty}^{\infty} f(x)dx = S$, em que S é um número real

positivo (não necessariamente igual a um), então $f(x)$ não satisfaz todas as condições para ser uma função densidade de probabilidade. No entanto, pode-se facilmente definir uma nova função g , em termos de f ,

por $g(x) = \frac{f(x)}{S}$ para todo x . Dessa maneira, $g(x)$ satisfará todas as

condições de uma função densidade de probabilidade (MEYER, 1983).

Segundo Ross (2000), a função densidade de probabilidade mais desenvolvida e utilizada é a normal ou Gaussiana, sendo utilizada para descrever o comportamento de uma variável aleatória que flutua de forma simétrica em torno de um valor central. Além disso, a distribuição normal está na origem de toda a formulação teórica acerca da construção de intervalos de confiança, testes estatísticos de hipóteses, bem como da teoria de regressão e correlação (TRIOLA, 2008).

A distribuição normal é um modelo com dois parâmetros, cuja função densidade é expressa pela Equação (9):

$$f(x, \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \cdot \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right] \text{ para } -\infty < x < \infty \quad (9)$$

onde μ e σ representam a média e o desvio padrão da distribuição, respectivamente.

No entanto, apesar das inúmeras qualidades inerentes à distribuição normal, pesquisas revelaram que em alguns casos os dados distribuem-se de forma assimétrica, caso não contemplado por essa notável distribuição (Weber, 2011). Para estes casos, existem outras funções densidade de probabilidade para representar tais distribuições como a Exponencial, Gama, *Weibull* e Lognormal.

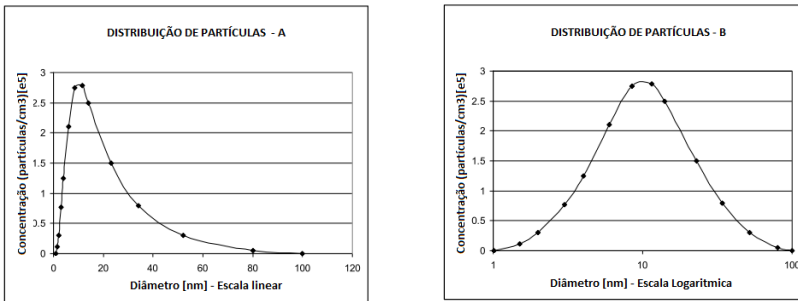
2.6.2 Distribuição Lognormal

Seja uma variável aleatória contínua X , definida em $0 \leq x < \infty$, resultante da ação multiplicativa de um grande número de componentes aleatórios independentes X_i ($i = 1, 2, \dots, n$), ou seja que $X = X_1 \cdot X_2 \dots X_n$. Nesse caso, a variável $Y = \log(X)$, tal que $Y = \log(X_1) + \log(X_2) + \dots + \log(X_n)$, em decorrência do teorema do limite central, irá tender a uma distribuição normal, com parâmetros μ_Y (média) e σ_Y (desvio padrão), quando n for suficientemente grande para permitir a convergência. Sob tais condições, diz-se que a variável X segue uma distribuição lognormal, com parâmetros $\mu_{\ln(X)}$ e $\sigma_{\ln(X)}$.

Em resumo, uma variável aleatória X terá distribuição lognormal se $\log(X)$ for normalmente distribuída, sendo usual a utilização do logaritmo natural $\ln(X)$. Além disso, o produto de variáveis aleatórias que seguem distribuição lognormal também seguirá uma distribuição lognormal (LIMPERT, STAHEL e ABBT, 2001).

No intuito de esclarecer como se diferencia uma distribuição lognormal, a Figura (9) traz uma representação de dados hipotéticos de distribuição de partículas onde os dados da distribuição, em escala linear, apresentam assimetria à direita, que é característica da distribuição lognormal (Figura 9a) e os dados da distribuição, já em escala logarítmica, apresentam-se simétricos, a qual é característica da distribuição normal (Figura 9b).

Figura 9 Dados de distribuição (a) em escala linear e (b) em escala logarítmica



Fonte: Adaptado de TSI (2013)

Assim, a distribuição lognormal também é um modelo com dois parâmetros, cuja função densidade lognormal é expressa pela Equação (10) (CROW e SHIMIZU, 1988):

$$f(x, \mu_{\ln(X)}, \sigma_{\ln(X)}) = \frac{1}{x\sigma_{\ln(X)}\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{\ln(x) - \mu_{\ln(X)}}{\sigma_{\ln(X)}}\right)^2\right], \text{ para } x > 0 \quad (10)$$

onde, x é a variável independente. Uma boa alternativa para estimar o parâmetro $\mu_{\ln(X)}$ é apresentada na Equação (11):

$$\mu_{\ln(X)} = \frac{\sum_{i=1}^n \ln(x_i)}{n} \quad (11)$$

No entanto, como a média aritmética dos logaritmos é igual ao logaritmo da média geométrica (Equação 12) a tendência central é utilizar a média geométrica $[\mu_g]$ no caso de uma distribuição lognormal.

$$\mu_{\ln(X)} = \frac{\sum_{i=1}^n \ln(x_i)}{n} = \frac{\ln x_1 + \ln x_2 + \dots + \ln x_n}{n} = \ln \sqrt[n]{x_1 \cdot x_2 \cdot \dots \cdot x_n} = \ln \mu_g \quad (12)$$

No caso do parâmetro $\sigma_{\ln(X)}$, uma opção adequada para estimá-lo é apresentada na Equação (13). Nesse caso, a medida de dispersão da distribuição lognormal é o desvio padrão geométrico $[\sigma_g]$, sendo que 95,4% dos valores pertencentes à distribuição estão dentro do intervalo expresso por $\ln \mu_g - 2 \ln \sigma_g < \ln x < \ln \mu_g + \ln \sigma_g$ (ZENDER, 2010).

$$\sigma_{\ln(X)} = \sqrt{\frac{\sum_{i=1}^n (\ln(x_i) - \ln \mu_g)^2}{n-1}} = \ln \sigma_g \quad (13)$$

De acordo com as Equações (12) e (13), a função densidade de probabilidade da distribuição lognormal pode ser definida pela relação da Equação (14) (CROW e SHIMIZU, 1988):

$$f(x, \mu_g, \sigma_g) = \frac{1}{x \ln \sigma_g \sqrt{2\pi}} \exp \left[-\frac{1}{2} \left(\frac{\ln(x / \mu_g)}{\ln \sigma_g} \right)^2 \right], \text{ para } x > 0 \quad (14)$$

No caso em que os dados estiverem agrupados em distribuições de frequências, os valores de x_i serão ponderados pelas respectivas frequências absolutas f_i , com $i = 1, 2, \dots, n$. Nesse caso, a média geométrica $[\mu_g]$ e o desvio padrão geométrico $[\sigma_g]$ podem ser definidos pelas relações apresentadas nas Equações (15) e (16), respectivamente (KLINK, PHILLIPS e DUNGAN, 2011):

$$\ln \mu_g = \frac{1}{N} \sum_{i=1}^n f_i \ln(x_i) = \ln \sqrt[n]{\prod_{i=1}^n (x_i)^{f_i}} \quad (15)$$

$$\ln \sigma_g = \sqrt{\frac{\sum_{i=1}^n f_i \cdot (\ln(x_i) - \ln \mu_g)^2}{N-1}} \quad (16)$$

com $N = \sum_{i=1}^n f_i$ sendo a frequência total absoluta.

No entanto, em situações reais, tais distribuições podem ser compostas por mais de uma moda, podendo ser bimodal, trimodal ou n-modal. Nestes casos, a função densidade de probabilidade pode ser

descrita por meio de uma soma ponderada de mais de uma função densidade de probabilidade lognormal unimodal (CROW e SHIMIZU, 1988). A Equação (17) descreve a função densidade de probabilidade lognormal multimodal:

$$FLM(x, \mu_g^i, \sigma_g^i, w^i) = \sum_{i=1}^n w^i \cdot f(x, \mu_g^i, \sigma_g^i) \quad (17)$$

onde $FLM(x, \sigma_g^i, \mu_g^i, w^i)$ é a função densidade de probabilidade lognormal multimodal; $f(x, \sigma_g^i, \mu_g^i)$ é a função densidade de probabilidade da moda i ; n é o número de modas e w^i é o fração da moda i admitindo que $\sum_{i=1}^n w^i = 1$.

2.6.3 Aplicações da Função Densidade Lognormal

Distribuições assimétricas são particularmente comuns quando as variações são grandes e os valores não podem ser negativos, como no caso da abundância de espécies, duração de período de latência de doenças infecciosas e a distribuição de recursos minerais na crosta da terra. Tais distribuições inclinadas, frequentemente são modeladas a partir de uma função densidade lognormal.

Na literatura são encontradas aplicações em diversas áreas que comprovam a eficiência e flexibilidade da referida função. Algumas áreas de aplicação e a forma de utilização são apresentadas a seguir:

- Engenharia de Materiais: modelagem das taxas de falha dos mecanismos de semicondutores, tal como corrosão, difusão, migração de metal e crescimento de fenda. Também é comumente empregada como uma distribuição para o tempo de reparo, bem como na análise das falhas por fadiga (RYAN, 2009; RAUSAND e HOYLAND, 2004).

- Tecnologia de Alimentos: caracterização de estruturas que representam o tamanho e frequência de partículas de gotas e ou bolhas geradas em processos de dispersão ou na passagem por poros no caso de membranas filtrantes (LIMPERT, STAHEL e ABBT, 2001). Destacam-se ainda trabalhos pertinentes a formação da distribuição do tamanho de cristais de gelo em processos de congelamento de alimentos (DONHOWE, HARTEL e BRADLEY JR, 1991) e ao desenvolvimento de um modelo estatístico da distribuição dos tempos de inativação microbiana em alimentos (PELEGA e COLEB, 1998).

- Geologia e Mineração: estimaco da concentrao de elementos qumicos na crosta terrestre e sua radioatividade geralmente seguem uma distribuico Lognormal (CLARCK e HARPER, 2010; SINGER, 2013).
- Cincias Ambientais: descrio do comportamento de distribuices de tamanho de partculas na formao de aerossis (RUITER e OESENBERG, 1987; TOBIAS, 1993; SEINFELD e PANDIS, 1997; MARING *et. al* 2003; YUAN *et. al* 2011), alm da modelagem de disperso de poluentes com o objetivo de realizar monitoramento da qualidade do ar. (OTTA, 1990; BURMASTER e HULL, 1997; TAYLOR, JAKEMAN e SIMPSON, 1986). A distribuico do tamanho de partcula de fuligem em um processo de combusto  representada com sucesso por uma funo Lognormal (JOHNSSON, BLADH e BENGTSSON, 2010; TIAN *et. al* 2004). Um excelente *review* sobre o tema pode ser encontrado em ZENDER (2010).
- Hidrologia simulaco da distribuico granulomtrica de solo e sua influncia na capacidade de reteno de gua como tambm na caracterizao da condutividade hidrulica saturada (CARSEL e PARRISH, 1988; TULI, KOSUGI e HOPMANS, 2001; FREDLUND, XING e HUANG, 1994; HWANG e CHOI, 2006).
- Economia: avaliao da probabilidade de sucesso de um investimento no mercado financeiro auxiliando na tomada de deciso (KLEBER e KOTZ, 2003; MYERS e MAJLUF, 1984).
- Ecologia: descrio da abundncia de espcies em comunidades de plantas e animais (MAGURRAN, 1988). Uma aplicao amplamente explorada pela engenharia florestal  a utilizao da funo lognormal para descrever distribuices diamtricas de rvores nativas ou plantadas (HAFLEY e SCHREUDER, 1977; CHOJNACKY, 1997).
- Medicina: descrio do perodo de incubaco de doenas infecciosas agudas (SARTWELL, 1950; HORNER, 1987; NISHIURA, 2007).
- Engenharia Qumica: inferncia da distribuico do tamanho de partcula atravs de RNs alimentadas por dados provenientes de um processo simulado de polimerizao de poliestireno (GUGLIOTTA *et. al* 2009). Estudos relacionados a microestruturas policristalinas representam a distribuico do tamanho de gros atravs da funo lognormal (GROEBER *et. al* 2008; ZHENG, ZHANG e CHEN, 2009).

2.7 CONSIDERAÇÕES SOBRE A REVISÃO BIBLIOGRÁFICA

Em síntese, de acordo com Lenzi (2002), a cinética da reação de polimerização em emulsão apresenta vantagens como a baixa viscosidade do meio reacional, a possibilidade de obtenção de polímeros de elevadas massas molares, facilidade de controle de temperatura, altas conversões e a possibilidade de aplicação direta do produto final.

No entanto, o controle de reatores de polimerização em emulsão é um problema extremamente complexo, devido a fatores como a falta de sensores *on line* para determinar propriedades específicas, como a DTP, o comportamento tipicamente não-linear e não-estacionário das reações e a grande complexidade da modelagem dinâmica (aspectos cinéticos e termodinâmicos em meios reacionais multifásicos) e da representação das relações estrutura/propriedades.

Apesar dos casos de sucesso relatados na literatura considerando a utilização das RNs no controle inferencial de processos de polimerização e da importância da DTP na determinação da qualidade do látex polimérico, não há trabalhos que utilizam RNs na inferência de DTP durante o processo de polimerização a partir de variáveis de operação.

Esta constatação se apresenta como um desafio, visto que não há trabalhos que corroborem a proposta, mas também como uma oportunidade. Isso porque a possibilidade de utilização de um conjunto de dados adequados e em diversas condições reacionais, associado a um algoritmo evolucionário eficiente como o PSO no treinamento de RNs MLP, pode se mostrar uma solução consistente na inferência de distribuição de tamanho de partículas nas reações de polimerização em emulsão.

3. MATERIAL E MÉTODOS

Este capítulo apresenta a procedência e algumas particularidades dos dados experimentais das reações de copolimerização em emulsão utilizados neste trabalho. Em seguida, discute a implementação de uma rede neural *feed-forward MLP* empregada para estimar a conversão e o diâmetro médio de partícula nestas reações, abordando aspectos relativos à arquitetura da rede (número de camadas ocultas e de neurônios por camada, função de ativação, bias), a utilização de dados atrasados no treinamento da rede e ao procedimento de treinamento da mesma. Os resultados da inferência da conversão e diâmetro médio de partícula são importantes para validar a estratégia proposta neste trabalho. O algoritmo de otimização global PSO, responsável pelo treinamento da rede neural (estimação dos pesos sinápticos da rede), tem também sua estratégia apresentada neste capítulo. O modelo baseado na função lognormal bimodal usado na inferência das DTPs também é apresentado nesse capítulo. Por fim, a estrutura da rede neural híbrida implementada para a inferência das DTPs é discutida, considerando aspectos relativos à arquitetura e utilização da conversão e Dp médio com entradas da rede, como também da utilização de uma estratégia de rede neural dinâmica em série, utilizando parâmetros da função lognormal bimodal, inicialmente determinados, como dados de entrada na inferência dos parâmetros seguintes.

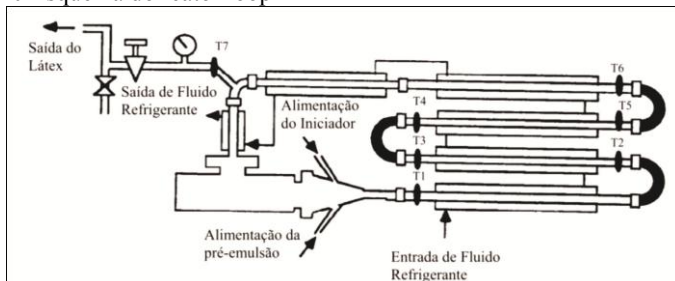
3.1 DADOS EXPERIMENTAIS

Os dados experimentais utilizados, como objeto de estudo, são oriundos do trabalho de Araújo (1999), tendo o autor investigado os efeitos das condições de operação sobre a distribuição do tamanho de partícula na copolimerização em emulsão em um reator *loop* contínuo de acetato de vinila e Veova10 com alto teor em sólidos. Um resumo do aparato experimental e sistema reacional são apresentados a seguir.

O reator *loop* contínuo utilizado, também conhecido como reator tubular contínuo com reciclo, está esquematizado na Figura (10) sendo este adaptado do trabalho de ABAD (1995). O reator é composto basicamente por uma tubulação encamisada, responsável pela troca térmica na reação; termopares que monitoram tanto a temperatura do reator quanto da camisa; bomba de recirculação; válvula de amostragem; e três bombas de alimentação dotadas de controladores mássicos, já que no caso de reatores contínuos, é necessário controlar cada uma das correntes independentemente, mantendo constante, desta forma, a relação entre as mesmas e a vazão global.

A alimentação do reator é feita perto da bomba de recirculação, enquanto parte do produto é retirado do reator pouco antes do ponto de alimentação, onde também é feito o controle da pressão para garantir que o reator permaneça sempre cheio.

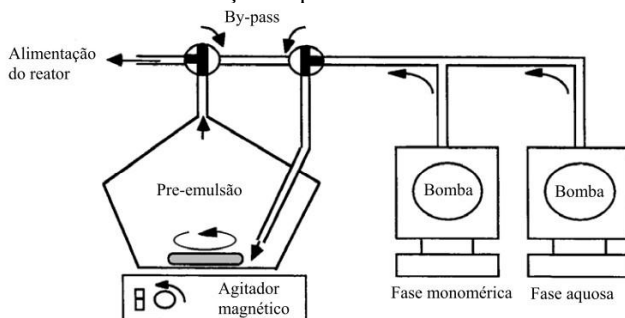
Figura 10 Esquema do reator loop



Fonte: Abad (1995)

Os reagentes são alimentados em três correntes diferentes: uma solução contendo o iniciador persulfato de potássio ($K_2S_2O_8$); uma fase aquosa contendo emulsificantes (Alipal (CO 436) e Arkopal (N-230)), colóide protetor hidroxietil celulose e o agente redutor metabissulfito de sódio ($Na_2S_2O_5$); a uma fase orgânica, constituída por uma mistura dos monômeros (acetato de vinila, veova10 e acrilato de butila, este último representando apenas 1 % em massa dos monômeros). De acordo com Abad *et. al* (1995), para o sistema de reação utilizado com teor de sólidos de 55% e usando um par redox como iniciador, a reação fica limitada por problemas difusionais. Para minimizar este problema, utilizou-se uma pré-emulsão com a fase aquosa e orgânica, como apresentado na Figura (11).

Figura 11 Sistema de alimentação da pré-emulsão



Fonte: Abad (1995)

Diante dessas considerações, a estratégia de partida do reator consistiu em encher o reator com a pré-emulsão estabilizando a temperatura de reação. Em todas as reações utilizou-se uma razão de recirculação elevada e igual a 55. Segundo Abad *et. al* (1995), este valor da razão de recirculação garante que o comportamento do fluxo deste reator loop se assemelhe ao de um reator contínuo de mistura perfeita (CSTR), aproximando-se das condições industriais de trabalho de um reator loop contínuo.

Quando a reação é iniciada, a partir da alimentação do iniciador, em função de uma concentração elevada de emulsificante e monômero no reator, o número de partículas inicialmente formado é muito elevado, o que dá origem a um brusco aumento da viscosidade, dificultando a estabilidade das taxas de alimentação. Com a diminuição do número de partículas devido à eluição e ao aumento do tamanho destas, a viscosidade também diminui até atingir um valor estacionário.

Esta estratégia de iniciação da reação propicia uma nucleação inicial muito forte devido à presença de um grande número de micelas. No decorrer da reação, a relativamente elevada solubilidade dos radicais oligoméricos na fase aquosa, por causa da presença do acetato de vinila, promove golpes de renucleação. Em alguns casos, estas renucleações permaneciam com uma determinada intensidade, produzindo fortes variações da DTP ao longo da reação.

Foram realizadas várias reações, com o intuito de estudar o efeito das variáveis de operação (temperatura da reação, tempo de residência, concentração de iniciador e concentração de emulsificante), sobre a dinâmica do reator loop e especialmente sobre a distribuição de tamanhos de partícula. A Tabela (4) sintetiza as condições de operação das reações investigadas por Araújo (1999).

Uma discussão sobre os efeitos das condições de operação sobre a conversão e diâmetro médio de partícula pode ser encontrada no Apêndice B, sendo importante para a escolha das variáveis de entrada da rede neural, bem como para a definição das reações que farão parte do treinamento e validação da mesma.

Além disso, dois procedimentos foram realizados a fim de adequar o conjunto de dados a este trabalho. O primeiro diz respeito à disponibilidade de DTPs experimentais e o segundo quanto à normalização das distribuições de frequência.

Tabela 4 Condições de operação das Reações de Polimerização realizadas por Araújo (1999)

Reação	R2	R4	R5	R7	R8	R9	R10
Tempo de residência (min)	13.3	13.3	13.3	13.3	13.3	13.3	13.3
Temperatura (°C)	60	50	50 / 60	60	60	60	60 / 50
Concentração de emulsificante (% em peso)	3.6	3.6	3.6	3.6 / 7.2	3.6	3.6	3.6
Concentração de iniciador (% em peso)	0.30	0.30	0.18	0.30	0.18	0.30/0.18	0.30

Fonte: Araújo (1999)

A determinação experimental das DTPs ao longo das reações realizadas por Araújo (1999) foi realizada através de Centrífuga Disco. Porém, em função do dispêndio de tempo necessário em cada análise, o autor somente realiza a determinação experimental da DTP em $\pm 25\%$ das amostragens. Por conta disso, o conjunto de dados experimentais é expandido considerando a interpolação linear das DTPs, sendo esta estratégia possível, visto que o crescimento do diâmetro das partículas poliméricas é linear no decorrer da reação. Uma discussão detalhada desta interpolação linear e a consequente expansão das distribuições experimentais está apresentada no Apêndice C.

O procedimento de normalização das distribuições de frequência absolutas se faz necessário para observar princípios relativos às distribuições de probabilidade. Assim, uma integração numérica foi realizada para o cálculo da área (S) sob a curva de distribuição de tamanho de partícula fazendo uso da regra do trapézio. A distribuição de frequências normalizadas (f_i^n) foi calculada a partir da distribuição de frequências absolutas (f_i), sendo definida pela Equação (18):

$$f_i^n = \frac{f_i}{S} \quad (18)$$

observando, nesse caso, que $\sum_{i=1}^k f_i^n = 1$ com $k = 1, \dots, 40$ é o número de agrupamentos por distribuição.

De posse das distribuições experimentais, bem como das distribuições construídas através da interpolação numérica foi possível gerar a superfície que descreve o comportamento da DTP durante toda a reação. O Apêndice D apresenta estas superfícies para as reações exploradas neste trabalho.

3.2 IMPLEMENTAÇÃO DO PSO

Nas últimas duas décadas, o número de trabalhos propondo melhorias na estratégia de implementação do PSO tem crescido exponencialmente, conforme discutido na seção 2.5. Observa-se, neste caso, que muitas das propostas avaliam a consistência da estratégia na minimização de funções teste que, nem sempre, possuem as adversidades presentes em um problema real, ou realizam esta avaliação na otimização de um problema muito específico, perdendo generalidade.

No entanto, algumas alterações se consolidaram como o uso do peso inercial. Destacando-se pela sua consistência, Jiao, Lian e Gu (2008), propõe uma estratégia de melhoria do PSO através da implementação do peso inercial dinâmico. A Equação (19) descreve o peso inercial dinâmico.

$$w = w_o \cdot u^{-k} \quad (19)$$

onde k é o número de iterações, w_o é o peso inercial inicial e u é um parâmetro de controle. Desse modo, permitem-se maiores variações na velocidade das partículas no início da busca, quando o valor de k é pequeno. Com a evolução do algoritmo, as soluções se aproximam do mínimo global e o aumento no valor de k possibilita uma busca mais refinada em torno desse valor.

No presente trabalho foram usados os seguintes parâmetros para o PSO: $u = 1.0008$, $w_o = 0.9$, $c_1 = c_2 = 1.4$, seguindo o proposto por Van den Bergh (2002) que investigou exaustivamente os critérios de convergência do algoritmo. Por fim, adotou-se um número de quarenta partículas por iteração, seguindo convenção apresentadas em artigos que investigam a eficiência do PSO.

O critério de inicialização está baseado em uma busca aleatória, a partir de uma distribuição uniforme, dentro do intervalo de busca

definido para cada parâmetro. Quanto aos critérios de parada, admite-se número máximo de 8000 iterações e a avaliação da Equação (20):

$$E = \sqrt{\sum_{n=1}^5 (P_{best} - P_n)} \quad (20)$$

onde P_{best} é o melhor resultado encontrado e P_n são os cinco últimos melhores resultados. Pelo critério, se $E < 10^{-5}$, a estimativa é finalizada.

3.3 DETERMINAÇÃO DA ARQUITETURA DA REDE NEURAL *MULTI-LAYER PERCEPTRON*

Dois pressupostos foram seguidos para fundamentar a estratégia de inferência por redes neurais neste trabalho. O primeiro diz respeito à escolha pela implementação de uma rede neural *Multi-Layer Perceptron*, sendo justificado pelo imenso histórico de sucessos relatados na literatura científica, pela sua flexibilidade em reconhecer inter-relações complexas entre variáveis, que não podem ser explicadas por modelos fenomenológicos, e pela sua baixa complexidade. O segundo está associado à utilização do algoritmo de otimização global *Particle Swarm Optimization* no ajuste dos pesos sinápticos da rede neural MLP, visto que trabalhos que utilizam sua heurística inteligente no treinamento de RNs se acumulam nos últimos anos.

RNs *Multi-layer Perceptron* foram exploradas nesse trabalho em dois momentos. Inicialmente, com o objetivo de validar a estratégia proposta na inferência da conversão e do diâmetro médio das partículas, conforme descrito na seção 3.4, sendo os resultados comparados com Araújo (2001a) que realizou estas mesmas inferências por meio de *Stacked Neural Network*. Também foram exploradas na determinação das DTPs em uma estrutura híbrida, sendo ela detalhada na seção 3.6 deste trabalho.

A literatura descreve a função tangente hiperbólica como uma das funções de ativação mais usuais, sendo utilizada em todas as redes implementadas e avaliadas nesse trabalho com estrutura de conexão *feed-forward*. Por sua vez, a topologia de uma rede MLP é formada pela camada de entrada, pela(s) camada(s) oculta(s), pela camada de saída e pela quantidade de neurônios que forma cada camada.

O dimensionamento da RN é um aspecto importante para o sucesso do seu treinamento, já que define o número de pesos sinápticos (parâmetros) que o PSO deve estimar. Considerando os vetores relativos às camadas que compõem a RN, o número de pesos a serem estimados no treinamento da RN MLP com uma única camada oculta pode ser calculado pela Equação (21):

$$N_{\text{pesos}} = [e + 1].[c] + [c + 1].s \quad (21)$$

onde, e = número de variáveis de entrada, c = número de neurônios na camada oculta e s = número de saídas da rede.

Segundo Basheer e Hajmeer (2000) o processo de configuração da arquitetura de uma rede MLP, número de camadas ocultas e número de neurônios na(s) camada(s) ocultas, é reportado por vários autores como um processo de tentativa e erro. Seguindo este processo, o Apêndice G apresenta o treinamento da rede neural MLP, que foi realizado em triplicata, com pesos iniciais diferentes, sendo estes gerados randomicamente pelo PSO. Os resultados dos erros de treinamento e validação para cada arquitetura testada, bem como o valor do desvio padrão em cada caso foram importantes na definição da arquitetura ideal, evitando problemas como *overfitting* e desperdício de tempo computacional.

3.4 DESENVOLVIMENTO DO MODELO NEURAL PARA PREDIÇÃO DE CONVERSÃO E DIÂMETRO MÉDIO DE PARTÍCULA

A inferência da conversão e Dp médio foram realizadas por RNs individuais, ou seja, uma RN foi treinada para a predição da conversão e outra RN foi utilizada para a predição do Dp médio. Com base nas condições operacionais mensuradas e controladas por Araújo (1999), ambas as RNs implementadas possuem como variáveis de entrada o tempo de residência – TR, a temperatura do reator - T, a temperatura da camisa - Tc, a fração do *holdup* - h , a concentração de iniciador $[i]$ e a concentração de surfactante $[s]$. A conversão $[X]$ e o diâmetro médio das partículas são as variáveis de saída de cada rede, respectivamente.

No sentido de oferecer a ambas as RNs mais subsídios para compreender as oscilações que ocorrem nas reações investigadas, muito em função de na reação R5 não ter sido controlada a temperatura do meio reacional, mas sim a temperatura da camisa do reator, foram usados dados atrasados da temperatura do reator e da camisa nas simulações. Nesse caso assumem também o papel de variáveis de entrada: $T_{[t-1]}$, $T_{[t-2]}$, $Tc_{[t-1]}$ e $Tc_{[t-2]}$, com um tempo de atraso de 10min para $[t-1]$ e 20min para $[t-2]$, aproximadamente, sendo por isso denominada Rede Neural Dinâmica (RND).

Em ambos os casos, uma segunda estratégia foi testada considerando também dados atrasados de conversão $X_{[t-1]}$ e $X_{[t-2]}$, para estimar a conversão atual, e dados atrasados de Dp médio $Dp_{[t-1]}$ e $Dp_{[t-2]}$ para estimar o Dp médio atual. Estas considerações são possíveis por causa da evolução de técnicas e metodologias de determinação de

conversão e Dp médio, permitindo a determinação dos mesmos em tempo real.

Nesta etapa do trabalho, os dados provenientes das reações R2, R4, R7 e R8, totalizando 160 padrões, foram utilizados para o treinamento da RN. Os dados provenientes das reações R5 e R9, totalizando 79 padrões, foram usados nesse trabalho para realizar a validação da rede neural, justamente para propiciar a comparação com Araújo *et. al* (2001b) já que os autores fazem uso destas reações na validação da RN proposta.

O ajuste dos pesos sinápticos de RNs pode ser formulado como um problema de otimização numérica, onde cada solução mantém um conjunto de valores para os pesos das conexões da rede e a função objetivo representa a soma dos erros cometidos pela rede para o conjunto de treinamento. Dessa forma, para o treinamento realizado, a função objetivo foi definida pela soma dos erros quadráticos normalizados (NSSE), definido pela Equação (22):

$$NSSE = \frac{1}{2} \sum_{n=1}^N \sum_{k=1}^C \left(t_k^n - o_k^n \right)^2 \quad (22)$$

onde, N representa o número de padrões do conjunto de dados de treinamento; C representa o número de unidades na camada de saída; t_k^n representa a saída normalizada desejada para a unidade de saída k para o padrão n e o_k^n representa a saída normalizada obtida pela rede na unidade k da camada de saída para o padrão n .

No intuito de permitir essa comparação entre resultados e, de certa forma, também seguir uma tendência de trabalhos que tratam de reconhecimento de padrões através de RNs, foi adotado o erro percentual médio, definido pela Equação (23), como parâmetro de controle da qualidade da estimação.

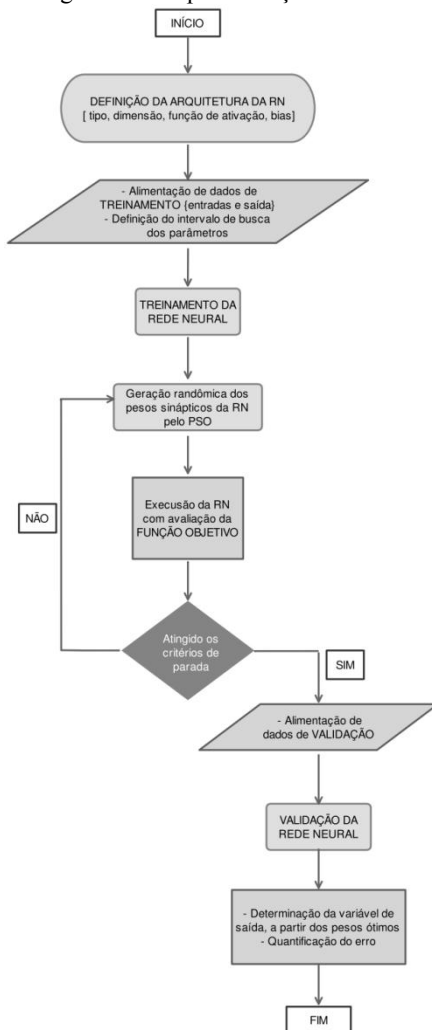
$$Erro[\%] = \sum_{i=1}^n \left| \frac{x_{exp}^i - x_{cal}^i}{x_{exp}^i} \right| \cdot \frac{100}{n} \quad (23)$$

onde n = número de pontos experimentais, x_{exp}^i = valor experimental e x_{cal}^i = valor predito pela rede neural dinâmica.

O treinamento da RND foi realizado em triplicata, de forma que os valores relatados para a função objetivo são a média aritmética dos resultados encontrados nos três processos. São relatados também os desvios padrão da função objetivo, sendo este parâmetro estatístico importante para mensurar a capacidade de reprodutibilidade da rede neural. Por fim, o coeficiente de correlação entre dados experimentais e preditos também é utilizado para avaliar a precisão da estratégia.

A Figura (12) apresenta um fluxograma do modelo neural implementado neste trabalho no intuito de realizar a inferência da conversão e do Dp médio. Cabendo salientar que o programa foi estruturado em linguagem *Fortran 90*.

Figura 12 Fluxograma de implementação da Rede Neural Dinâmica



Fonte: Desenvolvido pelo autor

3.5 MODELAGEM DA DISTRIBUIÇÃO DE TAMANHO DE PARTÍCULA USANDO FUNÇÃO LOGNORMAL MULTIMODAL

Desde o início do trabalho ficou claro que a inferência das distribuições de frequência absoluta (f_i) era inviável, visto que não há uma relação direta das frequências absolutas individuais de uma distribuição com as variáveis de operação, mas sim da curva de DTP com estas variáveis, sendo esta relação, inclusive, discutida por Araújo (1999). Assim, um dos objetivos fundamentais do trabalho foi encontrar uma função que pudesse descrever as DTPs com precisão e com um número reduzido de parâmetros.

A princípio, a distribuição de *Schulz-Flory* foi proposta para descrever estas distribuições, já que trabalhos a utilizaram com sucesso na construção de distribuições de massas moleculares (NELE e PINTO, 2002; FORTUNY *et. al* 2004). Uma adaptação da referida função apresentou desempenho satisfatório na descrição de distribuições unimodais, no entanto, apresentou inconsistências matemáticas quando escrita como uma soma ponderada de distribuições para descrever multimodalidades, sendo assim descartada.

Diante das características assimétricas das DTPs exploradas e dos inúmeros casos de sucesso em diversas áreas de conhecimento como ciências atmosféricas e ciências naturais, a função densidade de probabilidades lognormal se apresentou como uma alternativa viável para descrever tais distribuições. Valendo-se desse conhecimento preexistente e de sua flexibilidade, é que a função densidade de probabilidade lognormal multimodal foi proposta para descrever as DTPs nas reações de polimerização em emulsão.

A Equação (24) descreve a função densidade de probabilidade lognormal multimodal utilizada neste trabalho:

$$FLM(d, \mu_g^i, \sigma_g^i, w^i) = \sum_{i=1}^n \frac{w_i}{d \cdot \ln \sigma_g^i \sqrt{2\pi}} \cdot \exp \left[-\frac{1}{2} \left(\frac{\ln(d) - \ln(\mu_g^i)}{\ln \sigma_g^i} \right)^2 \right], \text{ para } d > 0 \quad (24)$$

onde $FLM(d, \sigma_g^i, \mu_g^i, w^i)$: Função Lognormal Multimodal; d é o diâmetro das partículas; n : número de modas e w^i : fração da moda i admitindo que $\sum_{i=1}^n w^i = 1$. Ressaltando que há três parâmetros a serem estimados em cada moda ($\ln \mu_g^i$: logaritmo natural da média geométrica

da moda i ; $\ln \sigma_g^i$: logaritmo natural do desvio padrão geométrico da moda i ; w^i : fração da moda i).

Diante das características das DTPs ao longo das reações investigadas neste trabalho, optou-se por realizar modelagens das distribuições fazendo uso de uma Função Lognormal Trimodal (FLT) e uma Função Lognormal Bimodal (FLB), comparando-se os resultados posteriormente.

A robustez do processo de estimação de parâmetros do modelo é garantida pelo *Particle Swarm Optimization* que usa sua estratégia inteligente para, de forma aleatória, minimizar a função objetivo definida pela soma dos erros quadrados (SSE), de acordo com a Equação (25):

$$SSE = \sum_{k=1}^n \left(f_{\text{exp}}^k - FLM(d_k) \right)^2 \quad (25)$$

onde f_{exp}^k : valor experimental da frequência no ponto k ; $FLM(d_k)$: frequência calculada pela FLM no ponto k ; n : número de pontos experimentais, d_k : diâmetro no ponto k .

Como o processo de estimação dos parâmetros realiza busca aleatória, é necessário definir os limites superiores e inferiores dos parâmetros a serem ajustados. O parâmetro “média geométrica” foi definido com base no domínio da variável independente “diâmetro” em escala logarítmica. Já os limites do parâmetro “desvio padrão geométrico” foi baseado na experiência de trabalhos que realizam a modelagem de DTPs usando função lognormal (LIMPET, STAHEL e ABBT, 2001). A Tabela (5) apresenta os intervalos de busca explorados, sendo que a estimativa inicial dos parâmetros é gerada aleatoriamente a partir de uma distribuição uniforme pelo PSO.

Tabela 5 Intervalo de busca dos parâmetros da FLM

Parâmetro	Intervalo de Busca
$\ln \mu_g$	[3.8, 6.6]
$\ln \sigma_g$	[0.01, 0.6]
w	[0, 1]

Fonte: Desenvolvido pelo autor

As regressões foram realizadas em triplicata, de forma que os valores relatados como parâmetros ótimos são a média aritmética dos

parâmetros encontrados nas três estimações. São relatados também os desvios padrão de cada parâmetro, salientando que estes desvios representam uma variabilidade do algoritmo em estimar aquele parâmetro específico já que o mesmo executa busca aleatória, não tendo relação com o intervalo de confiança do parâmetro ajustado. Ressalta-se ainda que a implementação da estratégia de estimação de parâmetros ocorreu a partir de um programa implementado em linguagem *Fortran 90*.

A interpretação estatística dos resultados foi realizada por meio do cálculo do coeficiente de correlação (r) entre as frequências experimentais e frequências calculadas pelo modelo após o ajuste dos parâmetros. Além disso, o teste de aderência de *Kolmogorov-Smirnov* foi utilizado para comparar as distribuições de frequência acumulada experimental com as distribuições de frequência acumulada estimada pela função lognormal multimodal. A metodologia que fundamenta o Teste de *Kolmogorov-Smirnov* é discutida no Apêndice F.

3.6 DESENVOLVIMENTO DO MODELO NEURAL HÍBRIDO PARA INFERÊNCIA DAS DTPs

Para a inferência das curvas de distribuição, um modelo híbrido foi desenvolvido. A estrutura deste modelo foi concebida utilizando a habilidade de representação das curvas de distribuição das funções lognormal bimodal, avaliada no presente trabalho, e a capacidade preditiva das RNs MLP.

Um passo importante neste processo foi considerar a conversão e o Dp médio como variáveis de entrada da rede neural híbrida. Esta decisão está pautada em duas premissas: a primeira é a de que a RND treinada com o algoritmo PSO pode inferir com sucesso a conversão e o diâmetro médio das partículas em uma reação de polimerização em emulsão, como demonstrado neste mesmo trabalho; a segunda se baseia na evolução significativa em relação a equipamentos e técnicas de determinação *on line* de variáveis como a conversão e o Dp médio, com destaque para a Espectroscopia de Infravermelho Próximo (NIRs).

Portanto em uma primeira abordagem, a partir de uma avaliação da importância das possíveis variáveis de entrada da rede neural para a evolução das DTPs, foram consideradas como variáveis de entrada o tempo de residência (TR) a temperatura do reator (T), a conversão (x), o diâmetro médio das partículas (Dp médio) e dados atrasados do diâmetro médio da partícula $Dp_{[t-1]}$, $Dp_{[t-2]}$, com um tempo de atraso de 10min para $[t-1]$ e 20min para $[t-2]$, aproximadamente. Vale salientar

que nesta estratégia inicial, a RND predizia simultaneamente todos os parâmetros da função lognormal bimodal.

A explicação para a supressão de variáveis como a concentração de surfactante, concentração de iniciador e fração do *holdup* em comparação com o modelo neural desenvolvido para prever X e DP médio está fundamentada na hipótese de que informações relevantes destas variáveis estão implícitas na variável conversão, agora admitida como variável de entrada da RND. A exclusão da variável temperatura da camisa (T_c) foi motivada pelo fato de que as condições de controle dessa variável na reação R5 são diferentes da maioria das reações usadas no treinamento.

Com relação à inclusão de dados atrasados de DP médio, estes têm o objetivo de conferir mais robustez à rede neural, justamente porque a dinâmica de evolução das DTPs está bastante relacionada com a evolução do DP médio no decorrer da reação.

A estratégia conta também com uma única camada oculta com dez neurônios, como discutido no Apêndice G, seguindo os mesmos pressupostos adotados na escolha da arquitetura da RND que realizou a inferência da conversão e DP médio.

Nesse trabalho, os parâmetros ajustados da FLB, que descrevem as curvas de DTPs experimentais provenientes das reações R2, R4, R7 e R8, R9 e R10, totalizando 160 padrões, foram utilizados para o treinamento das RNDs. Os parâmetros ajustados da FLB que descrevem as curvas de DTPs experimentais provenientes da reação R5, totalizando 28 padrões, foram utilizados para realizar a validação das RNDs. A escolha das distribuições da reação R5 para compor o conjunto de dados de validação está associado a duas situações: não existem variações nas condições reacionais em R5 que não estejam contempladas nas outras reações, sendo importante visto que se trata de estratégia que faz uso de redes neurais. Além disso, os valores limítrofes dos parâmetros ótimos da FLB para a reação R5 não extrapolam os valores limítrofes dos parâmetros ótimos da FLB para as demais reações, visto que RNs não realizam extrapolação com segurança. Vale salientar ainda que durante a reações R5 ocorrem uma série de renucleações que resultam em distribuições complexas, aumentando o desafio de prever tais DTPs.

Uma segunda estratégia foi denominada Rede Neural Dinâmica em Série [RNDS] também foi implementada, nesse caso, realizando a inferência dos parâmetros da função lognormal de forma individualizada. Este modelo de rede neural híbrido apresenta vantagens como à diminuição significativa do número de pesos sinápticos, já que as saídas são individualizadas e permite que parâmetros da FLB já

ajustados possam servir de entrada na inferência dos parâmetros restantes, caracterizando esta como uma rede neural dinâmica em série.

A fim de agilizar a representação dos parâmetros da FLB como saídas da RNDS, estes receberam uma codificação, de acordo com a Tabela (6):

Tabela 6 Codificação dos parâmetros da função lognormal bimodal

Parâmetro	FLB	Codificação
Média Geométrica da 1ª Moda	$\ln \mu_g^1$	M1
Média Geométrica da 2ª Moda	$\ln \mu_g^2$	M2
Desvio Padrão Geométrico da 1ª Moda	$\ln \sigma_g^1$	D1
Desvio Padrão Geométrico da 1ª Moda	$\ln \sigma_g^2$	D2
Fração da 1ª Moda	w^1	P1
Fração da 1ª Moda	w^2	P2

Fonte: Desenvolvido pelo autor

Como há uma sequência lógica a ser seguida nas inferências, as entradas (E) e saídas (S) foram numeradas para facilitar a discussão dos resultados posteriormente.

A Figura (13) apresenta um fluxograma do procedimento de inferência de cada parâmetro na função lognormal. Conforme a mesma, a inferência dos parâmetros da FLB foi iniciada pelo parâmetro M1, sendo este considerado a saída (S1) da rede neural e tendo como entradas somente variáveis do processo e dados atrasados de diâmetro médio (TR, T, (x), Dp, Dp_[t-1] e Dp_[t-2]), totalizando 6 variáveis de entrada (E1) Após o ajuste dos pesos sinápticos e consequente validação da RN, este parâmetro M1, média geométrica da primeira moda da DTP, passou a integrar o conjunto de dados de entrada da rede neural (E2), tendo o parâmetro M2 como saída RN (S2). Com o ajuste dos pesos sinápticos e consequente validação dessa nova estrutura, o processo segue com a inferência do parâmetro P2, sendo a nova saída (S3). Uma avaliação do comportamento dos parâmetros M1 e M2 indicou semelhanças quanto ao comportamento do parâmetro P2, sendo, portanto, adensadas às variáveis de entrada (E3). Como o parâmetro P2 representa a fração da segunda moda em relação à distribuição total, o parâmetro P1 foi definido por diferença, de acordo com a seguinte relação $P1=1-P2$. Com o conhecimento de quatro dos parâmetros da FLB, a inferência do parâmetro D1, desvio padrão geométrico da

primeira moda, classificado como saída (S4), aconteceu adotando M1, M2, P1 e P2 como variáveis de entrada juntamente com as variáveis de processo inicialmente descritas, sendo esta a entrada (E4). Para finalizar, após o ajuste dos pesos sinápticos e consequente validação da RN, é realizada a inferência do parâmetro D2, classificado como saída (S5). Nesse caso, todos os outros parâmetros da FLB integraram as variáveis de entrada, totalizando 11 variáveis de entrada (E5). De posse de todos os parâmetros da FLB inferidos pela rede neural dinâmica em série, foi possível construir as curvas de distribuição de tamanho de partícula comparando-as com as curvas de DTPs experimentais da reação de validação (R5).

Figura 13 Fluxograma identificando variáveis de entrada e saída da rede dinâmica em série para uma função lognormal bimodal



Como no caso da inferência de conversão e D_p médio, para o treinamento realizado a função objetivo foi definida pela soma dos erros quadráticos normalizados, Equação (22). Como o treinamento da rede neural foi realizado em triplicata, o desvio padrão também constitui ferramenta de controle da eficiência da estratégia, assim como o erro médio percentual, definido pela Equação (23).

Por fim, o modelo neural híbrido implementado para realizar a inferência dos parâmetros na FLB possui semelhanças com o modelo neural proposto na seção 3.4, inclusive com a mesma estrutura de implementação. Com o detalhe que neste caso, a rede neural não está inferindo propriedades diretamente relacionadas à reação [conversão e D_p médio], mas sim parâmetros que irão compor o modelo que descreve as DTPs em momentos distintos da reação

4. RESULTADOS E DISCUSSÕES

Este capítulo apresenta os resultados obtidos no desenvolvimento de um modelo neural para inferência da conversão e Dp médio. Também são apresentados os resultados referentes ao uso da função lognormal multimodal na modelagem das DTPs. Finalmente, são apresentados os resultados obtidos com o modelo neural híbrido para predição das DTPs.

4.1 INFERÊNCIA DA CONVERSÃO E DIÂMETRO MÉDIO

4.1.1 Arquitetura da Rede Neural MLP

O Apêndice G apresenta os resultados para os testes que definiram a arquitetura ideal da rede neural MLP implementada tanto para a determinação da conversão, quanto do Dp médio. Em síntese, uma única camada oculta apresentaria resultados similares ou superiores a redes MLP com duas ou mais camadas ocultas. Além disso, que dez neurônios na camada oculta são suficientes para que a rede neural reconheça um padrão entre dados de entrada e saída.

4.1.2 Inferência da Conversão

A tentativa inicial de estimar a conversão considerando como entrada da rede variáveis de processo e dados atrasados de temperatura não foi muito promissora. Os erros percentuais médios obtidos para a conversão na reação R5 e R9 foram 15,74% e 4,01%, respectivamente.

Por conta disso, dados atrasados de conversão foram integrados a entrada da RND a fim de conferir maior robustez à estratégia. A Tabela 4.1 apresenta os erros percentuais médios dos resultados da estimação da conversão nas reações R5 e R9 obtidos através do treinamento da rede, sendo que estes mostram a evolução da estratégia implementada em comparação com os resultados divulgados por Araújo *et. al* (2001b). Esta evolução pode ser creditada não somente a utilização de dados atrasados de conversão, mas também a utilização do algoritmo PSO no treinamento da RND.

Tabela 7 Erro percentual médio obtido nesse trabalho e por Araújo *et. al* (2001b), na predição da conversão para o grupo de validação

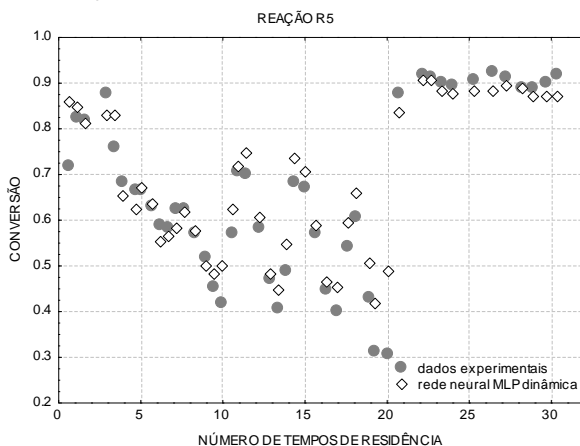
	Este trabalho	Araújo <i>et. al</i> (2001b)
Conversão - R5	7,33	17,26
Conversão - R9	3,43	14,28

Fonte: Desenvolvido pelo autor

As Figuras (14) e (15) apresentam a comparação entre dados experimentais de conversão e resultados da predição realizada pela RND para o grupo de validação. Em relação aos resultados da conversão para a reação R5, pode-se observar que mesmo na primeira metade da reação, onde ocorre um comportamento completamente oscilatório, não tendo sido atingido ainda o estado pseudo-estacionário, a estratégia proposta se mostrou sensível, reproduzindo com certa precisão essas fortes oscilações. Estas oscilações estão associadas à temperatura da reação que era de 50°C até $TR = 20,75$, atenuando-se a partir do momento, em que a temperatura da reação passou a ser de 60°C.

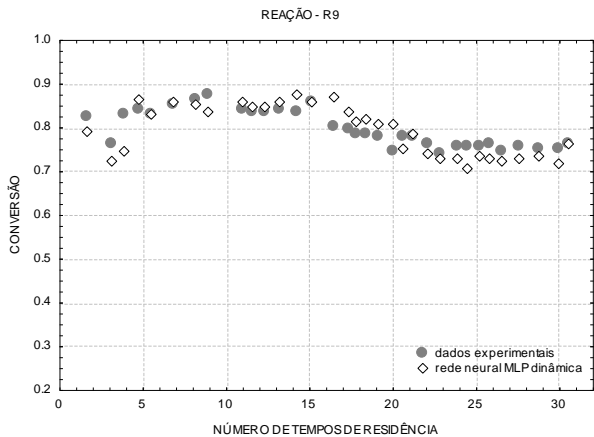
Quanto aos resultados da conversão para a reação R9 pode-se observar concordância entre dados experimentais e preditos pela rede neural, mesmo quando alterada a concentração de iniciador durante a reação, provocando o declínio da conversão.

Figura 14 Dados experimentais (ARAÚJO, 1999) e calculados pela RND para a conversão na reação R5



Fonte: Desenvolvido pelo autor

Figura 15 Dados experimentais (ARAÚJO,1999) e calculados pela RND para a conversão na reação R9



Fonte: Desenvolvido pelo autor

A capacidade adaptativa da RND na predição da conversão em ambos os casos pode ser reafirmada através da análise do coeficiente de correlação entre dados experimentais e preditos, como pode ser observado na Tabela (08).

Tabela 8 Coeficiente de correlação entre dados experimentais e preditos para a conversão pela RND

	RND com dados de conversão
Reação R5	0,916
Reação R9	0,969

Fonte: Desenvolvido pelo autor

4.1.3 Inferência do Diâmetro Médio de Partícula

A inferência do Dp médio a partir de variáveis de processo e dados atrasados de temperatura para as reações de validação R5 e R9 se mostraram promissoras. A Tabela (09) apresenta a comparação entre estes resultados e os obtidos por Araújo *et. al* (2001b) para as mesmas reações, imprimindo maior consistência à proposta.

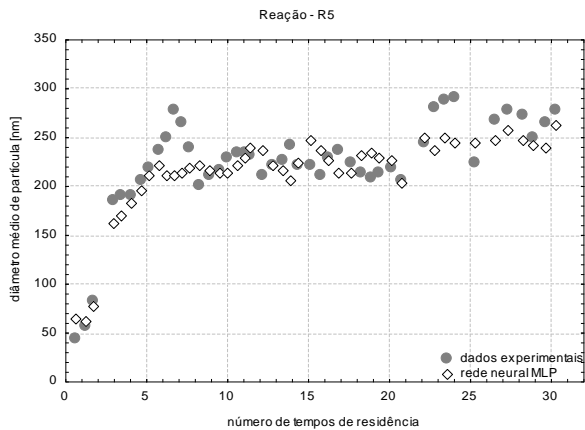
Tabela 9 Erro percentual médio obtido nesse trabalho e por Araújo *et. al* (2001b), na predição do Dp médio para o grupo de validação.

	Este trabalho	Araújo <i>et. al</i> (2001b)
Dp médio - R5	8,86	11,78
Dp médio - R9	4,71	7,89

Fonte: Desenvolvido pelo autor

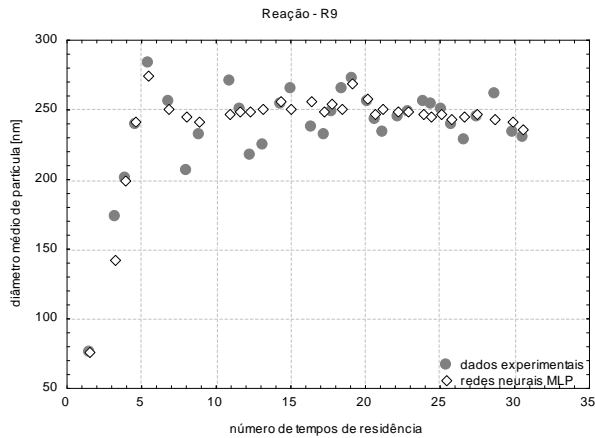
No entanto, apesar de a estratégia apresentar um erro percentual relativamente baixo em comparação com Araújo *et. al* (2001b), pode-se afirmar que ela não se mostrou tão sensível as oscilações no diâmetro médio de partícula, causadas por renucleações sucessivas, como pode ser observado nas Figuras (16) e (17). Possivelmente isso ocorreu porque as reações R5 e R9 possuem mudanças nas condições de processo durante as reações que não estão presentes em outras reações. Esta dificuldade foi superada na predição da conversão muito provavelmente porque nesse caso foram usados dados atrasados não só de temperatura que, segundo Araújo (1999), afetam consideravelmente a conversão, como também dados atrasados de conversão que forneceram a rede neural informações importantes sobre a dinâmica da reação.

Figura 16 Comparação entre dados experimentais (ARAÚJO, 1999) e calculados pela rede neural MLP para o diâmetro médio de partículas na reação R5



Fonte: Desenvolvido pelo autor

Figura 17 Comparação entre dados experimentais (ARAÚJO, 1999) e calculados pela rede neural MLP para o diâmetro médio de partículas na reação R9



Fonte: Desenvolvido pelo autor

Considerando o sucesso dos resultados alcançados na inferência da conversão, proposta similar foi testada para o caso do diâmetro médio de partícula, passando-se a considerar, como variável de entrada dados atrasados do diâmetro médio de partículas.

O erro percentual médio obtido na estimação do diâmetro médio de partícula através dessa rede neural dinâmica com dados atrasados de Dp médio foi de 6,94% para a reação R5, e 4,90% para a reação R9. Apesar de não haver uma diferença acentuada entre os valores do erro percentual em comparação com o erro no Dp médio sem dados atrasados de Dp médio apresentados na Tabela (09), observou-se que há uma diferença estatística relevante quando avaliado o coeficiente de correlação entre dados experimentais e calculados nas duas estimações, como pode ser observado nos resultados apresentados na Tabela (10).

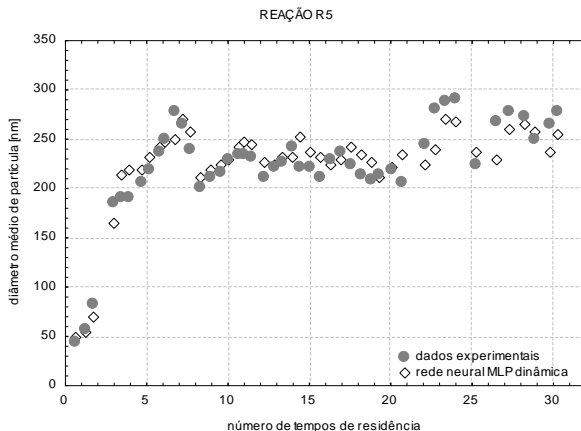
Tabela 10 Coeficiente de correlação entre dados experimentais e preditos para o Dp médio pela RND sem e com dados atrasados de Dp médio

	RND	RND com dados atrasados de Dp médio
Reação R5	0,901	0,936
Reação R9	0,916	0,924

Fonte: Desenvolvido pelo autor

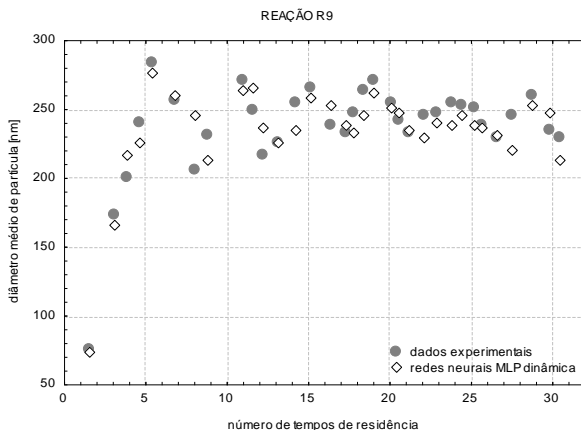
Esta constatação é corroborada pela observação das Figuras (18) e (19), mostrando que o uso da RND com dados atrasados de D_p médio foi mais sensível às oscilações do diâmetro médio de partícula.

Figura 18 Dados experimentais (ARAÚJO, 1999) e calculados pela RND com dados atrasados de D_p médio na inferência do D_p médio na reação R5



Fonte: Desenvolvido pelo autor

Figura 19 Dados experimentais (ARAÚJO, 1999) e calculados pela RND com dados atrasados de D_p médio na inferência do D_p médio na reação R9



Fonte: Desenvolvido pelo autor

Com base nos resultados apresentados observa-se que a RND, com pesos sinápticos estimados através do algoritmo estocástico PSO, é

capaz de representar a evolução da conversão e do diâmetro médio de partícula durante reações contínuas de polimerização em emulsão, mesmo nos casos em que se percebem muitas renucleações.

No entanto, é possível aumentar a eficiência e a precisão da RN com uma avaliação estatística mais criteriosa dos dados de treinamento e validação buscando apresentar à rede neural, através dos dados de treinamento, uma maior quantidade de oscilação existente nas reações para que esta possa captar melhor a complexidade das relações que definem a cinética de polimerização.

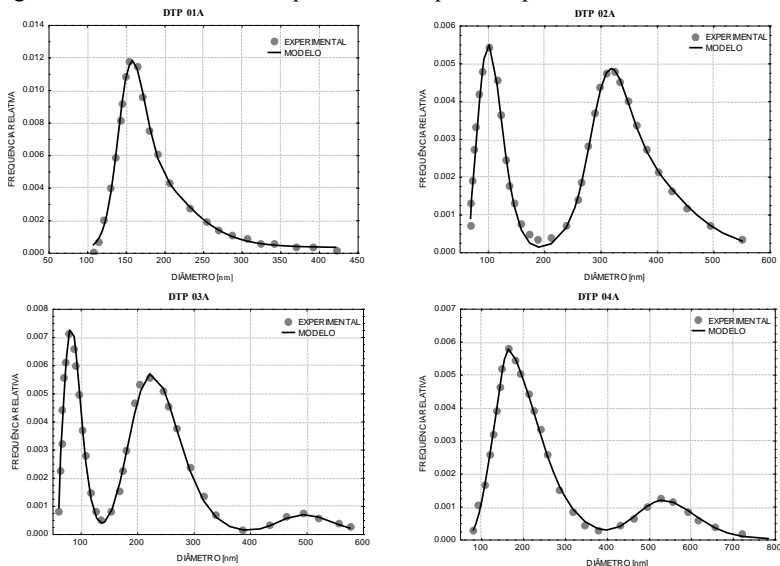
4.2 MODELAGEM DAS DISTRIBUIÇÕES DE TAMANHO DE PARTÍCULA USANDO A FUNÇÃO LOGNORMAL MULTIMODAL

A Função Lognormal Multimodal proposta pode construir curvas de DTP com mais de uma moda. A literatura científica apresenta casos de polimerização em emulsão onde aparecem essencialmente distribuições unimodais, bimodais e/ou trimodais. Portanto, a discussão a seguir está pautada na escolha entre uma Função Lognormal Bimodal ou Função Lognormal Trimodal para descrever as DTPs de polimerização em emulsão em reator *loop* contínuo, objeto de estudo desse trabalho.

Inicialmente, foram escolhidas DTPs experimentais típicas presentes na reação R5, considerando um caso em que a distribuição é unimodal (DTP 01), dois casos em que as distribuições são bimodais (DTPs 02 e 04) e um caso de distribuição trimodal (DTP 03). A discussão sobre as circunstâncias em que elas foram formadas é irrelevante neste momento, servindo, neste caso, somente para avaliar a eficiência da função lognormal bimodal e trimodal em descrever DTPs.

A comparação entre DTPs experimentais e modeladas pela FLT está apresentada na Figura (20). Os gráficos mostram claramente que a função lognormal trimodal se ajusta perfeitamente as distribuições experimentais, o que era de se esperar pelo número elevado de parâmetros que flexibiliza muito a construção de curvas de distribuição com três modas ou menos.

Figura 20 Curvas de DTPs experimentais e previstas pela FLT



Fonte: Desenvolvido pelo autor

A avaliação do coeficiente de correlação (r) entre distribuições de frequência experimentais e previstas é coerente com os resultados apresentados na Figura (20), visto que $r \geq 0.99$ para todos os casos onde se usou FLT, permitindo concluir que dados experimentais e previstos estão fortemente correlacionados. Com relação ao teste de aderência de *Kolmogorov-Smirnov*, o maior valor para $d_{\text{máx}}$ encontrado nos casos investigados foi $d_{\text{máx}} = 0.0045$. De acordo com os valores críticos estabelecidos pelo teste de *Kolmogorov-Smirnov*, deve-se admitir $d_{\text{crítico}} = 0.210$ para um nível de significância $\alpha = 5\%$ e para um número de pontos observados $n = 40$, aplicados neste trabalho. Assim, a hipótese de que os dados se referem a uma mesma distribuição deve ser aceita, já que $d_{\text{crítico}} = 0.210 > d_{\text{máx}} = 0.0045$ acrescentando credibilidade à modelagem realizada.

A Tabela (11) apresenta a média e o desvio padrão dos parâmetros ajustados usando FLT e o valor da função objetivo (SSE) ao final da regressão. O erro na estimação dos parâmetros foi quantificado a partir da análise da função objetivo, resultando em valores baixos em todos os casos investigados contribuindo para comprovar a robustez do PSO. Já os baixos desvios padrão dos parâmetros e função objetivo contribuíram para atestar a capacidade do algoritmo em reproduzir resultados de forma consistente.

Tabela 11 Parâmetros das curvas de DTP estimados usando FLT

Parâmetro \pm desvio padrão

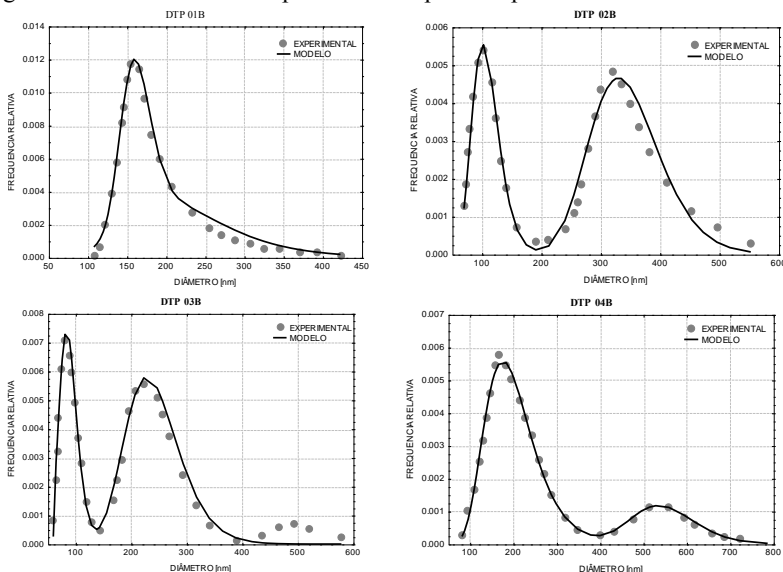
Reações	$\ln \mu_1$	$\ln \mu_2$	$\ln \mu_3$	$\ln \sigma_1$	$\ln \sigma_2$	$\ln \sigma_3$	w_1	w_2	w_3	SSE
DTP 01A	5.0673 \pm 0.0004	5.2588 \pm 0.0031	6.1347 \pm 0.0226	0.1061 \pm 0.0006	0.2378 \pm 0.0015	0.5051 \pm 0.0088	0.3346 \pm 0.0042	0.4939 \pm 0.0007	0.1716 \pm 0.0046	1.46E-06 \pm 3.42E-09
DTP 02A	5.1009 \pm 0.0008	5.2782 \pm 0.0004	6.3031 \pm 0.0003	0.0899 \pm 0.0007	0.3158 \pm 0.0004	0.1296 \pm 0.0004	0.0258 \pm 0.0005	0.7679 \pm 0.0001	0.2063 \pm 0.0005	2.93E-07 \pm 1.95E-10
DTP 03A	4.4635 \pm 0.0013	5.4603 \pm 0.0025	6.2168 \pm 0.0008	0.1837 \pm 0.0001	0.1904 \pm 0.0005	0.0963 \pm 0.0002	0.2896 \pm 0.0005	0.6288 \pm 0.0002	0.0815 \pm 0.0001	3.53E-06 \pm 1.95E-09
DTP 04A	4.6687 \pm 0.0002	5.7616 \pm 0.0034	5.9035 \pm 0.0030	0.2131 \pm 0.0002	0.0985 \pm 0.0002	0.2196 \pm 0.0033	0.3054 \pm 0.0001	0.2011 \pm 0.0118	0.4935 \pm 0.0119	6.61E-07 \pm 2.52E-10

Fonte: Desenvolvido pelo autor

Como um dos objetivos principais deste trabalho era encontrar uma função que pudesse descrever as curvas de distribuição de tamanho de partícula, através de uma estratégia inferencial usando RNs, foi necessário estar atento a condições que pudessem comprometer o treinamento da rede neural. A principal delas é o número excessivo de parâmetros da função lognormal multimodal utilizada que resultaria em um custo computacional muitas vezes desnecessário ou que inviabilizaria sua finalidade. Além disso, características relativas à nucleação e ao crescimento das partículas, e por consequência, a evolução das curvas de DTP ao longo das reações, podem não ser perfeitamente absorvidas pelos parâmetros estimados para cada curva. Isto porque, supõe-se que ao representar uma distribuição unimodal por uma FLT, pode ocorrer alta correlação entre os parâmetros devido ao grande número de parâmetros no modelo, como comprova a matriz de correlação apresentada no Apêndice E. Esta correlação pode atrapalhar a busca randômica realizada pelo PSO, uma vez que há a possibilidade de que o algoritmo encontre o mesmo resultado de ajuste aos dados experimentais com mais de um conjunto de parâmetros.

Estas constatações motivaram uma investigação sobre modelagem das DTPs usando a FLB. A Figura (21) apresenta a comparação entre distribuições experimentais e preditas usando FLB. As representações gráficas atestam a capacidade da referida função em predizer as DTPs, salientando que neste caso, considerou-se um número de parâmetros significativamente menor. Observa-se também que a FLB ignora a terceira moda na DTP 03B, visto que o modelo em questão só propicia a geração de duas modas, salientando que a moda ignorada é a menos representativa da distribuição.

Figura 21 Curvas de DTPs experimentais e preditas pela FLB



Fonte: Desenvolvido pelo autor

A avaliação estatística relativa à modelagem das DTPs usando FLB é muito semelhante à realizada quando se usa FLT. O coeficiente de correlação calculado indicou $r \geq 0,98$ para todos os casos investigados. Já para o teste de aderência de *Kolmogorov-Smirnov*, o maior valor para $d_{\text{máx}}$ encontrado nos casos investigados foi $d_{\text{máx}} = 0.0147$, significativamente menor que $d_{\text{crítico}} = 0.210$ e, portanto, admitindo a hipótese de que os dados se referem a uma mesma distribuição.

A Tabela (12) apresenta a média e o desvio padrão dos parâmetros ajustados da FLB e o valor da função objetivo (SSE) ao final da regressão. Os resultados da função objetivo mostram que não há perda significativa de eficiência na estratégia quando se usa a FLB. Além disso, os valores referentes aos desvios padrão dos parâmetros estimados e função objetivo são mínimos, contribuindo também para atestar a capacidade do algoritmo PSO em reproduzir resultados consistentes e ratificar a hipótese de que a redução do número de parâmetros faz com que o modelo absorva melhor a dinâmica de evolução das curvas de DTP durante as reações.

Um detalhe que aparentemente constitui-se em uma desvantagem na utilização da FLB é a existência de correlação entre seus parâmetros,

assim como no caso da FLT. O Apêndice E apresenta a matriz de correlação de parâmetros da FLB identificando estas correlações.

Uma análise da relação entre a evolução dos parâmetros da FLB e a dinâmica de evolução das DTPs ao longo das reações mostra que esta correlação é inevitável. Esta afirmação se baseia em algumas premissas, são elas: Os parâmetros $M1$ e $M2$ (média geométrica), que coincidem com as modas estatísticas da distribuição, evoluem no decorrer da reação em concordância com o crescimento das partículas, observando que nos casos em que há mais de uma moda estes parâmetros evoluem paralelamente. Esta correlação é ainda mais clara quando se observa o comportamento dos parâmetros $P1$ e $P2$ (fração da moda), visto que $P1 = 1 - P2$. Já as correlações que envolvem os parâmetros $D1$ e $D2$ são mais complexas, como pode ser observado pela própria Equação (24) que apresenta a FLM, salientando que no início das renucleações o parâmetro sofre um aumento, causado pelo crescimento rápido das partículas que produziu modas estreitas, atenuando-se no decorrer da reação acompanhando, nesta circunstância, a evolução do parâmetro média geométrica.

Por fim, a partir dos resultados obtidos, pode-se afirmar que a variabilidade presente nas distribuições de tamanho de partícula pode ser perfeitamente assimilada pela função lognormal bimodal, com um custo computacional menor que no caso da função lognormal trimodal, visto que no caso de FLB são cinco os parâmetros a serem estimados e não oito como no caso da FLT. Cabe ressaltar ainda que não foram observadas inconsistências no modelo, uma vez que foram observados os limites para o domínio de existência do diâmetro da partícula, o qual pode variar para diferentes mecanismos e sistemas reacionais, assim como considerar a utilização de frequências relativas normalizadas pela área da curva da distribuição.

Tabela 12 Parâmetros das curvas de DTP estimados usando FLB

Parâmetros \pm desvio padrão

Reação	$\ln \mu_1$	$\ln \mu_2$	$\ln \sigma_1$	$\ln \sigma_2$	w_1	w_2	SSE
DTP 01B	5.0830 \pm	5.3823 \pm	0.1178 \pm	0.3204 \pm	0.4354 \pm	0.5646 \pm	6.51E-06 \pm
	3.54E-08	2.73E-07	5.12E-08	9.42E-08	3.38E-07	3.38E-07	4.82E-11
DTP 02B	4.4658 \pm	5.4700 \pm	0.1861 \pm	0.2069 \pm	0.2963 \pm	0.7036 \pm	3.34E-06 \pm
	3.81E-08	1.73E-07	1.28E-07	6.16E-07	1.69E-07	1.69E-07	1.85E-11
DTP 03B	4.6718 \pm	5.8315 \pm	0.2160 \pm	0.1756 \pm	0.3127 \pm	0.6873 \pm	6.36E-06 \pm
	1.58E-07	3.41E-07	1.41E-07	1.53E-07	3.80E-07	3.80E-07	4.16E-11
DTP 04B	5.2643 \pm	6.3033 \pm	0.3038 \pm	0.1354 \pm	0.7851 \pm	0.2149 \pm	7.01E-07 \pm
	8.19E-09	1.23E-08	1.55E-08	4.17E-08	2.57E-08	2.57E-08	4.21E-12

Fonte: Desenvolvido pelo autor

4.3 INFERÊNCIA DAS DISTRIBUIÇÕES DE TAMANHO DE PARTÍCULA POR REDES NEURAIS DINÂMICAS EM SÉRIE

4.3.1 Inferência dos Parâmetros da Função lognormal Bimodal

Como discutido anteriormente, uma rede neural MLP demanda uma definição prévia da arquitetura ótima, que possa se adaptar com mais precisão ao problema de reconhecimento de padrões proposto. Nesse caso, como uma rede neural MLP já havia sido usada para prever conversão e D_p médio, optou-se por partir desta mesma configuração para avaliar qual arquitetura se mostrava mais efetiva da predição dos parâmetros da FLB.

No Apêndice G é realizada uma investigação a cerca da arquitetura ótima, ficando determinado que uma rede neural MLP com uma única camada oculta, composta por dez neurônios obteve resultados mais promissores e, portanto, adotada neste trabalho para realizar a inferência dos parâmetros da FLB.

Com uma estratégia que se mostrou inconsistente, a RND inicialmente proposta tinha como objetivo inferir simultaneamente os cinco parâmetros da FLB. Apesar da coerência dos dados experimentais, não havia diversidade em relação às variáveis de processo nas reações investigadas e o número de pesos sinápticos a serem ajustados era muito elevado, levando a estratégia a um desempenho nada satisfatório.

Posteriormente, optou-se por uma estratégia que realizava a inferência individual de cada parâmetro da FLB apresentando resultados promissores. Este capítulo discute os resultados da inferência dos parâmetros da FLB através de RND e RDNS.

A Tabela (13) apresenta uma comparação entre os erros médios percentuais na inferência dos parâmetros da FLB por RND e por RDNS. O erro médio percentual apresentado é relativo exclusivamente aos dados do conjunto de validação, totalizando 28 curvas da reação R5. De forma geral, os maiores erros estão associados ao parâmetro desvio padrão geométricos (D_1 e D_2). A explicação para estes resultados pode estar na distinta evolução do comportamento desse parâmetro no decorrer da reação, já que o mesmo assume valores maiores quando a distribuição está mais concentrada, referente ao início das renucleações, e valores menores quando a distribuição está mais dispersa.

Os resultados de M1 apresentaram as menores variações quanto ao erro médio percentual quando se compara RND e RDNS, isto porque, para este primeiro parâmetro, como discutido anteriormente, nenhum parâmetro foi adensado às variáveis de entrada para o treinamento com RDNS. A redução do erro médio percentual nesse caso pode ser

creditada a uma readequação pontual realizada em algumas curvas de DTPs pertencentes aos dados de treinamento e validação e que apresentavam resultados para o ajuste dos parâmetros da FLB pouco coerentes com a evolução das demais curvas de DTP pertencentes à mesma reação. Considera-se que esta readequação seria insuficiente para causar mudanças significativas na inferência por RND, mas tão somente ajustes e por conta disso não foram realizadas para a referida rede.

Tabela 13 Erros médios percentuais da RND e da RNDS

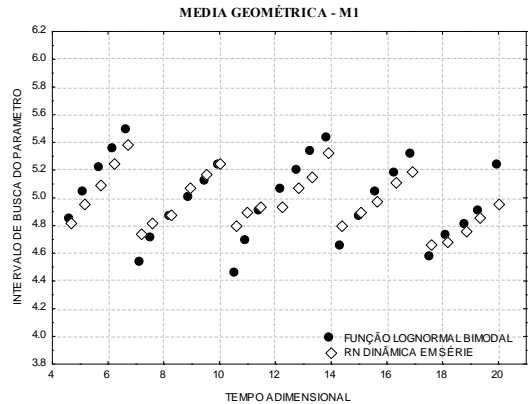
Parâmetro	Erro RND	Erro RNDS
M1	2,84%	2,07%
M2	2,67%	1,71%
D1	14,77%	8,51%
D2	27,61%	16,04%
P1	12,06%	2,43%
P2	15,26%	2,88%

Fonte: Desenvolvido pelo autor

Em resumo, como indicam os resultados da Tabela 4.7, a estratégia que utiliza os parâmetros previamente inferidos pela RN como variável de entrada imprime informações importantes a RN.

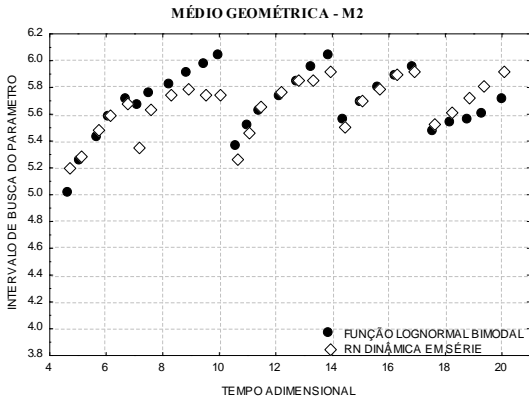
As Figuras (22) a (27) apresentam um comparativo entre parâmetros ajustados da FLB e os parâmetros inferidos pela RNDS para o conjunto de dados da validação (reação R5), salientando que os parâmetros M1, M2 e P2 têm comportamento semelhante à variável de entrada - diâmetro médio da partícula (D_p) que cresce linearmente no decorrer da reação até ocorrer a nucleação de novas partículas, sendo que o reconhecimento dessa conexão facilitou o processo de inferência dos parâmetros citados e justifica a utilização de dados atrasados de D_p .

Figura 22 Parâmetro M1 ajustado da FLB e predito pela RNDS



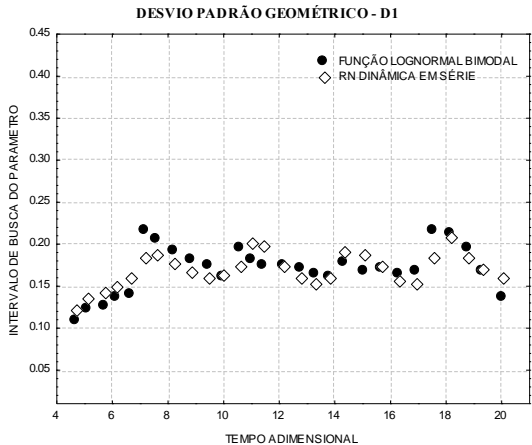
Fonte: Desenvolvido pelo autor

Figura 23 Parâmetro M2 ajustado da FLB e predito pela RNDS



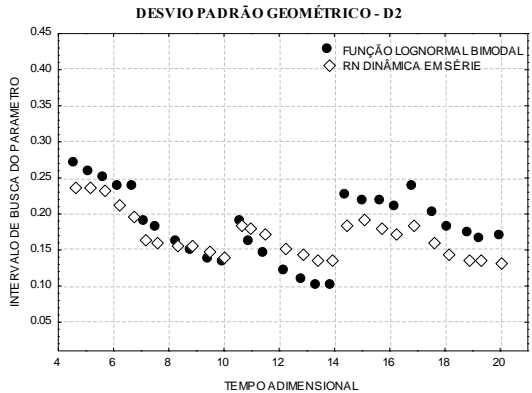
Fonte: Desenvolvido pelo autor

Figura 24 Parâmetro D1 ajustado da FLB e predito pela RNDS



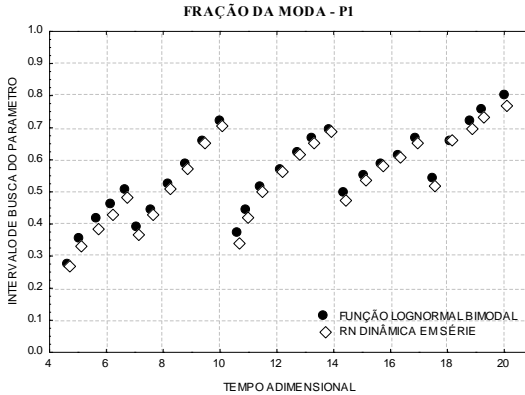
Fonte: Desenvolvido pelo autor

Figura 25 Parâmetro D2 ajustado da FLB e predito pela RNDS



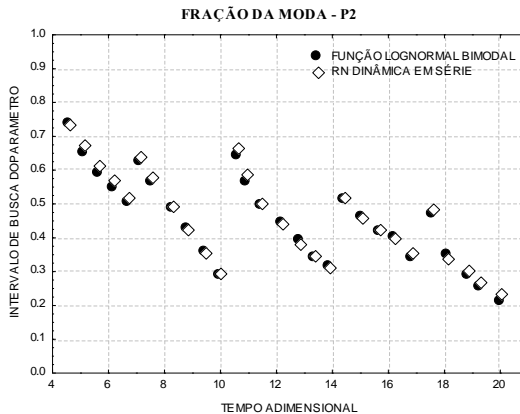
Fonte: Desenvolvido pelo autor

Figura 26 Parâmetro P1 ajustado da FLB e predito pela RNDS



Fonte: Desenvolvido pelo autor

Figura 27 Parâmetro P2 ajustado da FLB e predito pela RNDS



Fonte: Desenvolvido pelo autor

Os valores dos parâmetros e desvios padrão da FLB, resultado da inferência por RNDS, são apresentados na Tabela (14). Também estão presentes na mesma tabela os resultados da minimização da função objetivo [SSE] para cada DTP predita. Os valores da função objetivo evidenciam que a RN tem mais dificuldades na inferência das DTPs que precedem as renucleações, o que é compreensível, visto que no momento em que surge uma nova moda ocorre uma alteração sensível nas distribuições. O fenômeno da renucleação pode ser percebido na Tabela (14) pela redução acentuada dos parâmetros M1 e M2, ou seja, nos tempos de residência iguais a 7,2; 10,6; 14,4 e 17,6.

Tabela 14 Parâmetros e desvios padrão da FLB obtidos na inferência por RNDs

TR	M1	M2	D1	D2	P1	P2	SSE
4.7	4.82 ± 0.04	5.19 ± 0.03	0.1198 ± 0.0123	0.2360 ± 0.0148	0.2641 ± 0.0046	0.7359 ± 0.0046	2.92E-04
5.1	4.96 ± 0.11	5.29 ± 0.05	0.1348 ± 0.0110	0.2365 ± 0.0123	0.3306 ± 0.0203	0.6694 ± 0.0203	5.56E-05
5.7	5.10 ± 0.01	5.49 ± 0.08	0.1434 ± 0.0151	0.2312 ± 0.0115	0.3868 ± 0.0130	0.6132 ± 0.0130	1.11E-04
6.2	5.25 ± 0.03	5.59 ± 0.06	0.1483 ± 0.0165	0.2098 ± 0.0206	0.4301 ± 0.0112	0.5699 ± 0.0112	6.29E-05
6.7	5.39 ± 0.02	5.67 ± 0.09	0.1593 ± 0.0087	0.1978 ± 0.0294	0.4796 ± 0.0259	0.5204 ± 0.0259	4.54E-05
7.2	4.74 ± 0.05	5.35 ± 0.01	0.1845 ± 0.0241	0.1625 ± 0.0117	0.3632 ± 0.0107	0.6368 ± 0.0107	4.61E-04
7.6	4.81 ± 0.06	5.63 ± 0.02	0.1872 ± 0.0265	0.1580 ± 0.0119	0.4250 ± 0.0140	0.5750 ± 0.0140	1.08E-04
8.3	4.87 ± 0.07	5.75 ± 0.03	0.1756 ± 0.0161	0.1568 ± 0.0199	0.5059 ± 0.0068	0.4941 ± 0.0068	1.77E-05
8.9	5.07 ± 0.05	5.79 ± 0.01	0.1657 ± 0.0088	0.1562 ± 0.0166	0.5740 ± 0.0158	0.4260 ± 0.0158	8.08E-05
9.5	5.17 ± 0.01	5.70 ± 0.06	0.1591 ± 0.0023	0.1489 ± 0.0121	0.6493 ± 0.0216	0.3507 ± 0.0216	9.59E-05
10.0	5.24 ± 0.02	5.74 ± 0.04	0.1619 ± 0.0043	0.1410 ± 0.0072	0.7084 ± 0.0229	0.2916 ± 0.0229	4.08E-05
10.6	4.80 ± 0.12	5.27 ± 0.12	0.1725 ± 0.0091	0.1697 ± 0.0423	0.3382 ± 0.0104	0.6618 ± 0.0104	5.36E-04
11.0	4.90 ± 0.11	5.47 ± 0.04	0.1996 ± 0.0295	0.1816 ± 0.0102	0.4158 ± 0.0233	0.5842 ± 0.0233	2.42E-04
11.5	4.93 ± 0.11	5.66 ± 0.03	0.1958 ± 0.0238	0.1711 ± 0.0044	0.5002 ± 0.0257	0.4998 ± 0.0257	4.18E-05
12.2	4.94 ± 0.04	5.77 ± 0.11	0.1747 ± 0.0066	0.1500 ± 0.0077	0.5634 ± 0.0170	0.4366 ± 0.0170	1.83E-04
12.8	5.07 ± 0.02	5.85 ± 0.07	0.1608 ± 0.0097	0.1417 ± 0.0083	0.6167 ± 0.0183	0.3833 ± 0.0183	2.27E-04
13.4	5.15 ± 0.05	5.85 ± 0.07	0.1520 ± 0.0082	0.1367 ± 0.0133	0.6545 ± 0.0217	0.3455 ± 0.0217	3.23E-04
13.9	5.32 ± 0.05	5.93 ± 0.09	0.1901 ± 0.0078	0.1352 ± 0.0063	0.6854 ± 0.0062	0.3146 ± 0.0062	9.04E-05
14.4	4.80 ± 0.16	5.50 ± 0.05	0.2167 ± 0.0195	0.1761 ± 0.0236	0.4778 ± 0.0339	0.5222 ± 0.0339	1.49E-04
15.1	4.90 ± 0.09	5.71 ± 0.04	0.1869 ± 0.0094	0.1923 ± 0.0116	0.5406 ± 0.0261	0.4594 ± 0.0261	3.65E-05
15.7	4.98 ± 0.09	5.78 ± 0.10	0.1717 ± 0.0088	0.1813 ± 0.0167	0.5781 ± 0.0082	0.4219 ± 0.0082	3.28E-05
16.3	5.11 ± 0.05	5.89 ± 0.06	0.1543 ± 0.0126	0.1722 ± 0.0211	0.6056 ± 0.0171	0.3944 ± 0.0171	5.60E-05
16.9	5.18 ± 0.04	5.93 ± 0.05	0.1508 ± 0.0118	0.1823 ± 0.0178	0.6496 ± 0.0450	0.3504 ± 0.0450	1.93E-04

17.6	4.66 ± 0.07	5.53 ± 0.03	0.1848 ± 0.0181	0.1613 ± 0.0243	0.5204 ± 0.0106	0.4796 ± 0.0106	1.34E-04
18.2	4.68 ± 0.09	5.61 ± 0.07	0.2065 ± 0.0283	0.1452 ± 0.0214	0.6634 ± 0.0137	0.3366 ± 0.0137	5.36E-05
18.9	4.77 ± 0.08	5.71 ± 0.16	0.1825 ± 0.0115	0.1360 ± 0.0222	0.7006 ± 0.0197	0.2994 ± 0.0197	5.55E-05
19.3	4.86 ± 0.09	5.81 ± 0.10	0.1698 ± 0.0108	0.1343 ± 0.0209	0.7288 ± 0.0184	0.2712 ± 0.0184	7.48E-05
20.1	4.95 ± 0.09	5.91 ± 0.09	0.1602 ± 0.0051	0.1324 ± 0.0153	0.7657 ± 0.0063	0.2343 ± 0.0063	1.30E-03

Fonte: Desenvolvido pelo autor

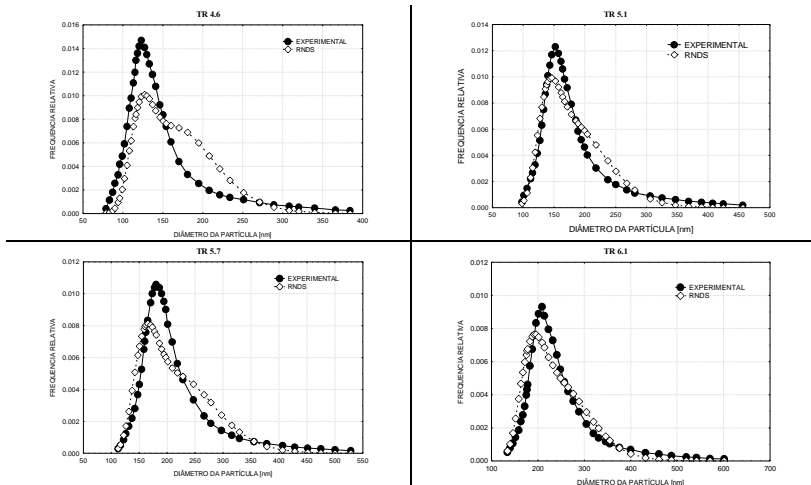
Outra constatação importante presente da Tabela (14) é relativa aos baixos valores dos desvios padrão, conferindo mais uma vez consistência à proposta de treinamento da rede neural com o PSO.

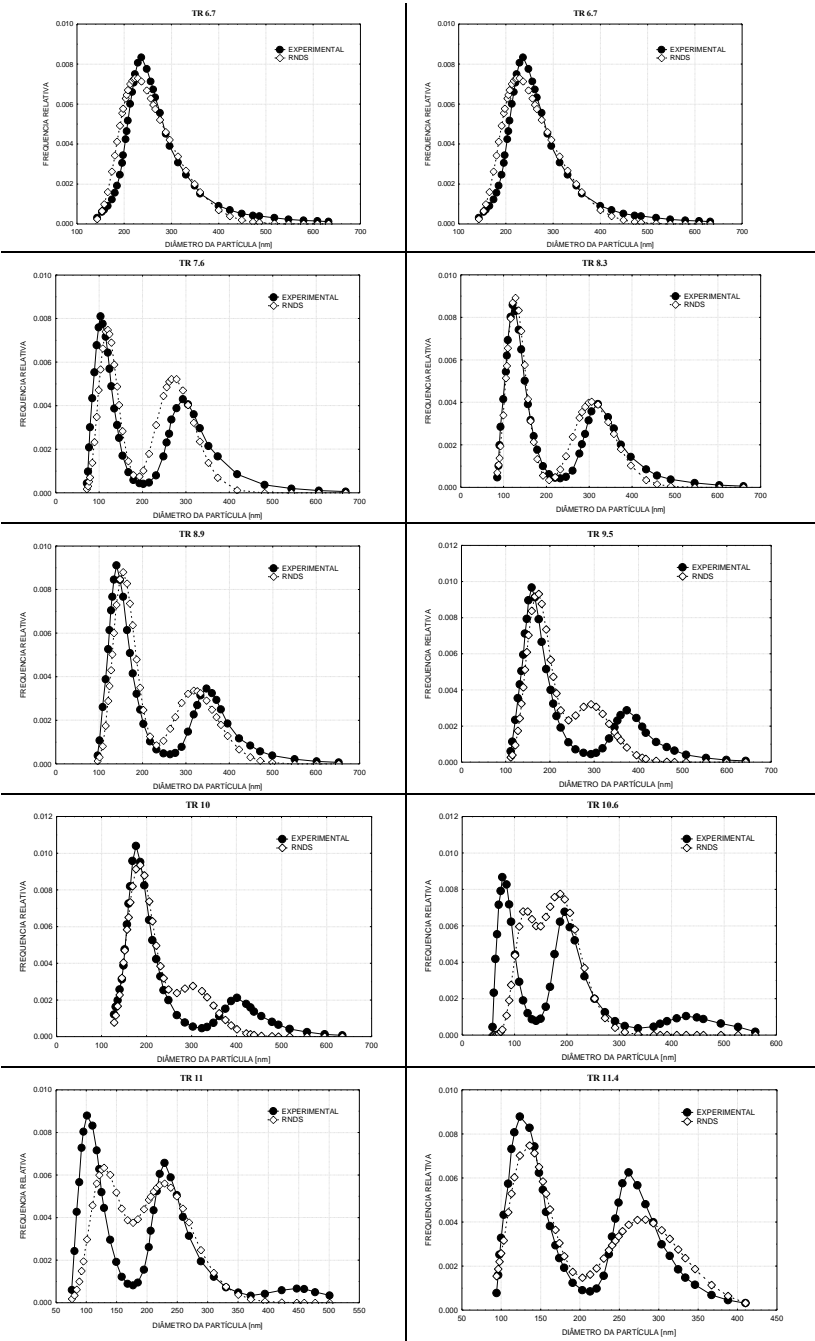
4.3.2 Construção das DTPs previstas

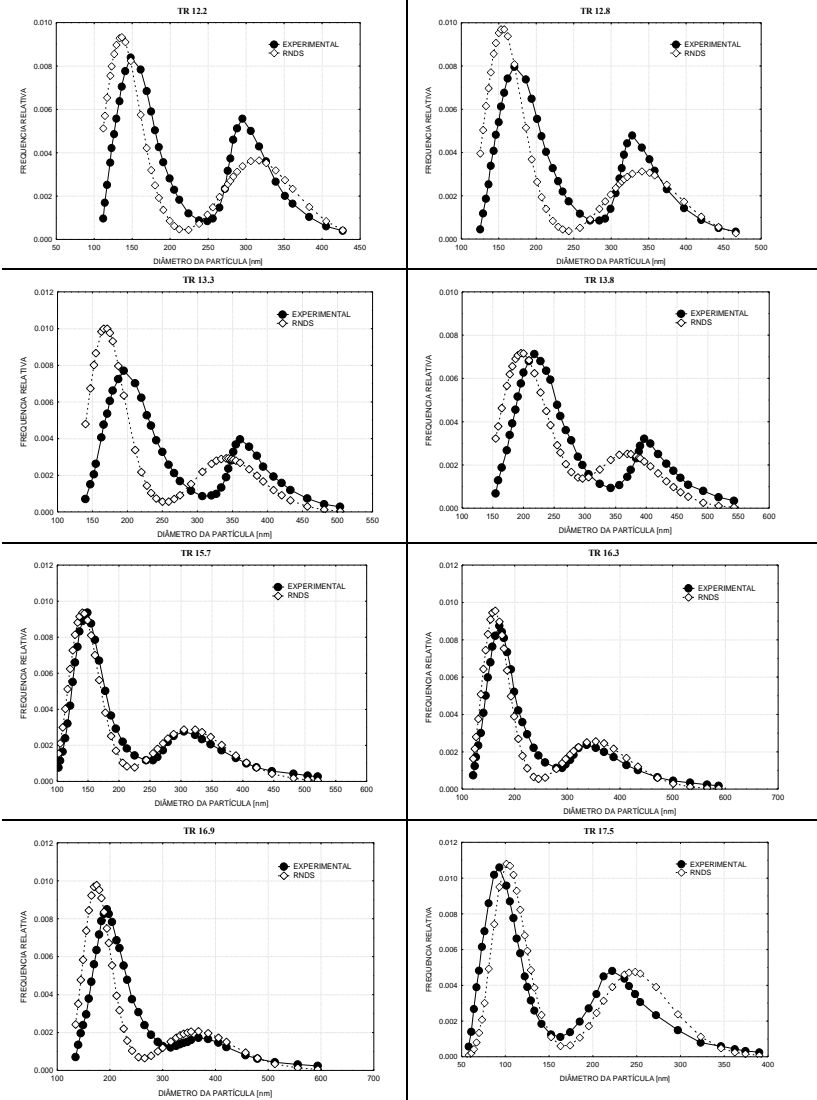
A partir da predição dos parâmetros da FLB foi possível construir as curvas de DTPs e realizar a comparação com as curvas de DTPs experimentais do conjunto de validação. A Figura (28) apresenta a comparação entre todas as distribuições experimentais e previstas para a reação de validação [R5], possibilitando assim, uma análise mais criteriosa da eficiência da rede implementada na tentativa de reproduzir a dinâmica de evolução das DTPs.

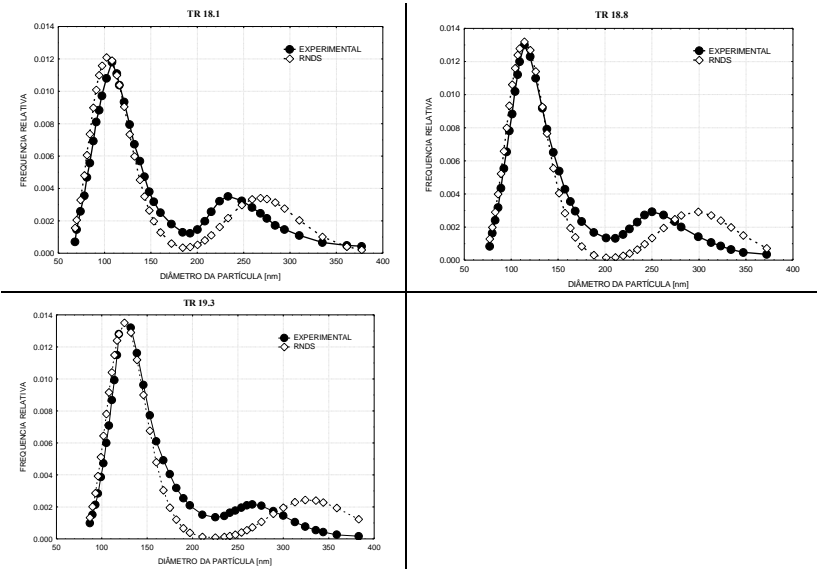
O modelo híbrido proposto apresenta algumas dificuldades em se adaptar as alterações bruscas nas características das distribuições, causadas após as renucleações. No entanto, nos momentos em que a dinâmica de evolução das distribuições se mostra mais estável ou mesmo em casos específicos de renucleações [TR = 17.5] a estratégia consegue prever com consistência e precisão as DTPs da reação.

Figura 28 Comparação entre curvas de DTPs experimentais e previstas da reação R5









Fonte: Desenvolvido pelo autor

A análise dos coeficientes de correlação entre as DTPs experimentais e preditas pela RNDs em diferentes momentos da reação, apresentados na Tabela (15), é mais um subsídio para afirmar a robustez da estratégia proposta. A exceção a este bom desempenho, esta em alguns resultados de DTPs preditas logo após as renucleações, como era de se presumir.

Tabela 15 Coeficientes de correlação entre DTPs experimentais típicas e inferidas pela RNDs

TR	Coefficiente de correlação (r)
5.7	0,95
6.2	0,97
8.3	0,93
10.0	0,92
11.0	0,74
12.2	0,54
14.4	0,93
16.9	0,72
18.2	0,97

Fonte: Desenvolvido pelo autor

Por fim, na observância do teste de *Kolmogorov-Smirnov*, o maior valor para $d_{\text{máx}}$ encontrado nos casos apresentados foi $d_{\text{máx}} = 0,0349$, consideravelmente menor que $d_{\text{crítico}} = 0,210$ e, portanto,

corroborando com a hipótese de que os dados se referem a uma mesma distribuição.

Diante dos resultados, é coerente afirmar que a estratégia de inferência por RNDS de distribuição de tamanho de partícula é promissora, e que esta pode ser uma alternativa coerente e precisa para prever DTP durante reações contínuas de polimerização em emulsão, servindo de ferramenta no monitoramento e controle de da qualidade do látex polimérico.

5 CONCLUSÕES E SUGESTÕES

5.1 CONCLUSÕES

O intuito inicial deste trabalho foi realizar a inferência das distribuições de tamanho de partícula em reações de polimerização em emulsão através de RNs treinadas com o auxílio do algoritmo *Particle Swarm Optimization*. Neste contexto, acredita-se que alguns fatores foram preponderantes para os resultados obtidos neste trabalho, sendo estes discutidos nos parágrafos a seguir.

A quantidade e qualidade dos dados experimentais, oriundos do trabalho de Araújo (1999) objeto de estudo deste trabalho, influenciaram positivamente os resultados. A determinação experimental sistemática das variáveis de operação durante as reações, a possibilidade de explorar dados em diversas condições reacionais e o conhecimento relativo ao crescimento linear dos diâmetros de partícula possibilitando a realização de interpolações que aumentaram significativamente o número de DTPs ao longo das reações, constituem as razões que corroboram com a afirmação inicial.

Em relação ao uso das RNs para prever conversão e D_p médio foi possível concluir que a estratégia que usou dados atrasados, formando assim uma rede neural dinâmica, foi superior em comparação com os resultados de Araújo *et. al* (2001b). O treinamento dessa rede através da estimação dos pesos sinápticos usando PSO se mostrou consistente, não sendo observados valores elevados para o desvio padrão da conversão e D_p médio preditos.

A consistência matemática da função densidade lognormal, fazendo uso de distribuições de frequências relativas normalizadas pela área de cada curva, associado ao conhecimento presente na literatura científica sobre o comportamento dos parâmetros da referida função na modelagem de diferentes perfis de distribuição de tamanho de partícula, também constitui fator importante para sucesso da estratégia. Verificou-se, pela comparação da FLT e FLB que a última foi mais adequada para representar as curvas de DTPs visto que o número de distribuições trimodais presentes no conjunto de dados de treinamento e validação era muito pequeno, sendo insuficiente para que a Rede Neural Dinâmica em Série [com inferência seqüencial dos parâmetros do modelo] as reconheça como um padrão diferenciado. Além disso, o uso de uma FLT permite ao algoritmo, que realiza busca estocástica, encontrar vetores de parâmetros ótimos diferentes para uma mesma distribuição, afetando a dinâmica de evolução dos vetores de parâmetros ótimos das DTPs no decorrer da reação e dificultando o trato destes dados por parte da rede

neural. Constatase ainda que o menor número de parâmetros da FLB reduziu significativamente o custo computacional. Por fim, é importante salientar que a predição das DTPs limitasse a faixa de condições reacionais investigadas, não sendo recomendado que se façam extrapolações.

Com relação ao desenvolvimento do modelo híbrido para predição das DTPs ressalta-se a simplicidade da rede neural implementada bem como sua robustez que está associada a detalhes como: a utilização de dados atrasados, transferindo à RN informações acerca da dinâmica da reação; o emprego de parâmetros da FLB previamente inferidos como variáveis de entrada, reforçando o treinamento da rede, visto que existe correlação entre estes parâmetros; e a consistência do algoritmo *Particle Swarm Optimization* que, através de uma heurística inteligente, realiza uma busca exaustiva dos pesos sinápticos —ótimos” no treinamento da RNDS.

Desconhecendo outras publicações que realizem a inferência por RNs das DTPs em reações contínuas de polimerização em emulsão, faz-se necessário ressaltar que a estratégia que se mostrou promissora pode ainda evoluir, principalmente na escolha das variáveis de operação que irão compor a entrada da rede neural bem como na diversidade das condições reacionais investigadas.

No entanto, vale ressaltar que, em reações contínuas de polimerização, podem ocorrer comportamentos oscilatórios devido à renucleação de partículas e à multiplicidade de estados estacionários para uma determinada faixa de condições de operação, tendo como consequência a formação de DTPs polidispersas. Logo, admite-se que para sistemas reacionais mais estáveis estas DTPs apresentem oscilações menores e o desempenho da estratégia proposta possa ser ainda melhor.

5.2 SUGESTÕES

Dentre as sugestões futuras indicadas para a continuidade deste trabalho estão:

Realizar um planejamento de experimentos que permita que os resultados possuam diversidade suficiente para que, na hipótese de implantação de uma rede neural para inferência de DTP, as informações mais relevantes a formação das distribuições estejam presentes.

Aplicar a estratégia de inferências de DTPs por RNs a outras técnicas, mecanismos e sistemas reacionais de polimerização a fim de verificar a sua adaptabilidade e eficiência.

Propor uma metodologia de pré-seleção das variáveis de operação da reação de polimerização, que seriam as entradas da rede neural para a inferência de DTPs, considerando não somente aspectos estatísticos, mas também a influência dessas variáveis sobre a DTP.

Investigar a implementação de outros tipos de RNs na inferência das DTPs como as Redes Neurais Recorrentes, que segundo a literatura científica apresentam facilidade de implementação, redução do número de neurônios e agilidade no processo de treinamento. Outros algoritmos de otimização global (*genetic algorithm, simulated annealing...*) também podem ser testados no treinamento da RN.

REFERÊNCIAS

ABAD, C. Estudio del Proceso de Obtención en un Reactor Loop de Copolímeros de Características Comerciales, Tese de Doutorado, Universidad del País Vasco. 1995.

ABAD, C; DE LA CAL, J. C; ASÚA, J. M. Modeling nucleation and particle growth in emulsion copolymerization in continuous loop reactors. *Macromol Symp.* 92, p.195–204, 1995.

AHMADI, M. A. Neural network based unified particle swarm optimization for prediction of asphaltene precipitation. *Fluid Phase Equilibria.* 314, 2012.

ALEXOPOULOS, A. H.; MAGGIORIS, C.; KIPARISSIDES, C. —CD Analysis of Turbulence Non-homogeneity in Mixing Vessels, a Two-compartment Model”. *Chemical Engineering Science*, vol. 57, pp. 1735-1752, 2002.

ALEXOPOULOS, A.H; MAGGIORIS, D.; KIPARISSIDES, C., —CD Analysis of turbulence non-homogeneity in Mixing Vessels a Two-compartment Model”, *Chem. Eng. Sci.*, vol. 57, pp. 1735-1752, 2002.

ALEXOPOULOS, A.H; ROUSSOS, A. I; KIPARISSIDES, C. Part I - Dynamic evolution of the particle size distribution in particulate processes undergoing combined particle growth and aggregation. *Chem Eng Sci.* 59, p.5751–69, 2004.

AMARAL, M.; VAN ES, S.; ASUA, J.M. Effect of the particle size distribution on the low shear viscosity of high-solid-content latexes. *Journal of Polymer Science. Part A: Polymer Chemistry.* 42, p.3936-3946, 2004.

ANDERSON, D. R.; SWEENEY, D. J.; WILLIAMS, T. A. Estatística aplicada à Administração e Economia. Trad. SANTOS, J. C. B. 2 ed. São Paulo: Editora CENGAGE Learning. 597 p. 2008.

ARAÚJO P. H. H., SAYER C., DE LA CAL J.C., ASUA J.M., LIMA E.L., PINTO J.C. Utilization of neural networks as soft sensors to monitor emulsion polymerization reactions (average particle diameter e conversion). *Latin American Applied Research*, 31:525-531. 2001b.

ARAUJO, P. H. H. ; DE LA CAL, J. C. ; ASUA, J. M. ; PINTO, J. C. Modeling Particle Size Distribution (PSD) in Emulsion

Copolymerization Reactions in a Continuous Loop Reactor. *Macromolecular Theory and Simulations*, v. 10, p. 769-779, 2001a.

ARAÚJO, P. H. H. Distribuição de Tamanho de Partícula em Sistemas Heterogêneos de Polimerização. Tese de Doutorado em Engenharia Química – Programa de Engenharia Química /COPPE – UFRJ, Rio de Janeiro, 1999.

ASSIS, A. J. Identificação e controle de processos não lineares utilizando RNs artificiais. Tese de Doutorado, Unicamp, Campinas, Brasil, 2001.

ASSIS, F. N. de; ARRUDA, H. V. de; PEREIRA, A. R. Aplicações de estatística à climatologia: teoria e prática. Pelotas: UFPEL, 1996.

BARBETA, P. A.; REIS, M. M.; BORNIA, A. C. Estatística para cursos de engenharia e informática. 2ª ed. São Paulo: Atlas, 2009.

BASHEER, I. A.; HAJMEER, M. Artificial neural networks: fundamentals, computing, design, and application. *Journal of Microbiological Methods*, v. 43, p. 3–31, 2000.

BLUM, C; MERKLE, D. *Swarm Intelligence – Introduction e Applications*. Springer-Verlag Berlin Heidelberg, 2008.

BOUTTI, S.; GRAILLAT, C.; McKENNA, T.F. High solids content emulsion polymerization without intermediate seeds. Part I. Concentrated monomodal lattices, *Polymer* 46, p. 1189-1201, 2005.

BRAGA, A. P.; CARVALHO, A. C. P. L. F.; LUDEMIR, T. B. RNs artificiais: teoria e aplicações. Rio de Janeiro: LTC, 2000.

BURDEN, R. L.; FAIRES, J. D. *Numerical Analysis*. CENGAGE Learning – São Paulo, 2008.

BURMASTER, D. E.; HULL, D. A. Using lognormal distributions and lognormal probability plots in probabilistic risk assessments. *Human and Ecological Risk Assessment: An International Journal* 3:2, 1997.

CAMPOS, H. de. Estatística experimental não-paramétrica. 4ª ed. Piracicaba: Departamento de Matemática e Estatística - ESALQ, 1983.

CAO L., LI D., ZHANG C., WU H. Control and modeling of temperature distribution in a tubular polymerization process. *Computers and Chemical Engineering*. Vol.31, p.1516–1524, 2007.

CAPEK, I; POTISK, P. Microemulsion and emulsion polymerization of butyl acrylate—I. Effect of the initiator type and temperature. *European Polymer Journal*. Vol. 31, Issue 12, p. 1269–1277, 1995.

CARSEL, R. F. PARRISH, R. S. Developing joint probability distributions of soil water retention characteristics. *Water Resources Research*. Vol. 24, Issue 5, 1988.

CARVALHO, M; LUDERMIR, T. B. Particle Swarm Optimization of Neural Network Architectures and Weights. *Seventh International Conference on Hybrid Intelligent Systems*. IEEE Computer Society, 2007.

CHAN, W. M; NASCIMENTO, C. A. O. Use of neural networks for modeling of olefin polymerization on high-pressure tubular reactors, *J. Appl. Poly. Sci.* Vol.53, p.1277–1289, 1994.

CHATERJEE, A.; SIARRY, P. Nonlinear Inertia Weight Variation for Dynamic Adaptation in Particle Swarm Optimization. *Computers e Operations Research*, 33, p. 859-871, 2004.

CHEN, Y; CHANG, K. C; HO, S. H. Improved framework for particle swarm optimization: Swarm intelligence with diversity-guided random walking. *Expert Systems with Applications*. Vol.38, 2011.

CHERN, C. S. Emulsion polymerization mechanisms and kinetics. *Progress in Polymer Science*. v.31, 443–486. 2006.

CHICOMA, D. L.; SAYER, C.; GIUDICI, R. In-Line Monitoring of Particle Size during Emulsion Polymerization under Different Operational Conditions using NIR Spectroscopy. *Macromol. React. Eng.* v.5, 150–162, 2011.

CHOJNACKY, D.C. Modeling diameter growth for pinyon and juniper trees in dry land forests. *Forest Ecology and Management*, vol. 93, Issues 1–2, 1997

CLARCK, I.; HARPER, W. V. *Practical Geostatistics*. Ecosse North America LLC, Columbus, Ohio, USA. 2000. Online: <http://faculty.otterbein.edu/wharper/pg2000forweb.pdf>, Updated: 12/09/2010.

COEN, E. M; PEACH, S; MORRISON, B. R; GILBERT, R. G; First-principles calculation of particle formation in emulsion polymerization: pseudo-bulk systems. *Polymer*, Oxford, v.45, n.11, p. 3595-3608, 2004.

COEN, E.M., GILBERT, R.G.; MORRISON, B.R.; LEUBE, H., PEACH, S. —“Modelling Particle Size Distributions and Secondary Particle Formation in Emulsion Polymerization”, *Polymer*, v.39, n.26, pp.7099-7112. 1998

CONSTANT, S. Modelagem de Reatores de Polimerização: Determinística e por RNs. Tese de doutorado apresentada a Faculdade de Engenharia Química da Universidade Estadual de Campinas – UNICAMP. (2007)

CROW, E.L., SHIMIZU, K. Log-normal Distributions: Theory and Application. New York: Dekker. 1988.

CROWLEY, T; MEADOWS, E; KOSTOULAS, E; DOYLE III, F. Control of particle size distribution described by a population balance model of semi batch emulsion polymerization. *J Proc Cont*, 10: 419–432, 2000.

CYBENKO, G. Approximation by superposition of a sigmoid function. *Mathematics of Control, Signals and Systems*, vol. 02, p. 303-314. 1989.

DE LANEY, D; OLIVER, S. On-line viscosity measurements on fiber forming polymers. In *Proceedings of the Society of Plastics Engineers—SPE 61st annual technical conference*. 2002.

DELFA, G. M; OLIVIERI, A; BOSCHETTI, C. E. Multiple response optimization of styrene–butadiene rubber emulsion polymerization. *Computers and Chemical Engineering*. Vol.33, p.850–856, 2009.

DONHOWE, D.P.; HARTEL, R.W.; BRADLEY JR, R.L. Determination of Ice Crystal Size Distributions in Frozen Desserts. *Journal of Dairy Science*. vol. 74, Issue 10, 1991.

EPKO, E; MUJTABA, I. M. Evaluation of neural networks-based controllers in batch polymerization of methyl methacrylate. *Neurocomputing*. Vol.71, p.1401–1412, 2008.

ETCHELLS, J; SNEE, T; WILDAY, J. Relief system design for exothermic runaway: the HSE strategy, *IChemE Symp. Ser. Vol.147*, p.299–311. 2000.

FERNANDES, F. A. N.; LONA, L. M. F. Development of Polymer Resins using Neural Networks, *Polímeros: Ciência e Tecnologia*, v. 12, n.3, p. 164-170, 2002.

FERNANDES, F. A. N; LONA, L. M. F. Neural Network Applications in Polymerization Processes. Brazilian Journal of Chemical Engineering. Vol. 22. Issue 03, p. 401-418, 2005.

FERNANDES, F. A. N; LONA, L. M. F; PENLIDIS, A. Inverse modeling applications in emulsion polymerization of vinyl acetate. Chemical Engineering Science, 59, 3159–3167, 2004.

FÉVOTTE, G., BARUDIO, I., GUILLOT, J., –An adaptive inferential measurement strategy for on-line monitoring of conversion in polymerization processes” Thermochimica Acta, v. 289, p. 223-242, 1996.

FORTUNY, M; NELE, M; PRIAMO, A. M; PINTO, J. C. Distribuição de Peso Molecular em Poliolefinas: Contribuição ao Estudo da Bimodalidade e da Desconvolução. Polímeros: Ciência e Tecnologia, vol. 14, nº 3, p. 194-200, 2004.

FREDLUND, D.G.; XING, A.; HUANG, S. Predicting the permeability function for unsaturated soils using the soil-water characteristic curve. Canadian Geotechnical Journal, vol. 31, Issue 4, 1994.

GAO, S. Z; WANG, J. S; GAO, X, W. Modeling and advanced control method of PVC polymerization process. Journal of Process Control 23 664–681; 2013.

GILBERT, R. G. Emulsion polymerization: A Mechanistic Approach. San Diego: Academic Press; 1995.

GIORDANI, D. S. –Estudo comparativo entre observador de estado não-linear de alto ganho e calorimetria associada a RNs artificiais para previsão de composição de copolímeros produzidos em emulsão com altos teores de sólidos”. Tese de Doutorado apresentada à Faculdade de Engenharia Química – UNICAMP. Campinas, São Paulo, 2003.

GROEBER, M; SOMNATH, G; UCHIC, M. D; DIMIDUK, D. M. A framework for automated analysis and simulation of 3D polycrystalline microstructures.: Part 1: Statistical characterization. Acta Materialia, Volume 56, Issue 6, April 2008.

GONZAGA, J. C. B; MELEIRO, L. A. C; KIANG, C; MACIEL FILHO, R. ANN-based soft-sensor for real-time process monitoring and control of an industrial polymerization process. Computers and Chemical Engineering. Vol.33, p.43–49, 2009.

GORI, M.; TESI, A. On the problem of local minima in back-propagation, IEEE Trans. Pattern Anal. Mach. Intell. Vol.14, Issue 1, p. 76–86, 1992.

GUGLIOTTA, L. M; STEGMAYER, G. S; CLEMENTI, L. A; GONZALEZ, V. D. G; MINARI, R. J; LEIZA, J. R; VEGA, J. R. A Neural Network Model for Estimating the Particle Size Distribution of Dilute Latex from Multiangle Dynamic Light Scattering Measurements. Part. Part. Syst. Charact. Vol.26, p.41–52, 2009.

GUYOT, A; CHU, F; SCHNEIDER, M; GRAILLAT, C; Mc KENNA, T. F. High solid content latexes. Progress in Polymer Science. Vol. 27, Issue 8, p. 1573–1615, 2002.

HAFLEY, W. L.; SCHREUDER, H. T. Statistical distributions for fitting diameter and height data in even-aged stands. Canadian Journal of Forest Research, 7(3), 1977.

HAGAN, M. T.; DEMUTH, H. B.; BEALE, M. Neural Network Design, 7th International Student Edition. Vikas Publishing House, 2003.

HANAI, T.; OHKI, T.; HONDA, H.; KOBAYASHI, T., Analysis of initial conditions for polymerization reaction using fuzzy neural network and genetic algorithm, Computers and Chemical Engineering, v. 27, p. 1011-1019, 2003.

HANDAN, S. A. Hybrid Particle Swarm Optimizer using multi-neighborhood topologies. Department of Computing and Information Technology Arab Open University Kuwait. 2008.

HAYKIN, Simon. Redes Neurais: princípios e prática. 2aed – Porto Alegre: Bookman, 2001.

HERRERA, F; ZHANG, J. Optimal control of batch processes using particle swam optimization with stacked neural network models. Computers and Chemical Engineering. Vol.33, p.1593–1601, 2009.

HORNER, R. D. Age at onset of Alzheimer's disease: clue to the relative importance of etiologic factors? American Journal Epidemiology, 126 (3), 1987.

HOSSEINI, A., BOUASWAIG, A., ENGELL, S. Comparison of classical population balance models of emulsion polymerization with

experimental results and a stochastic extension. Chem. Eng. Sci. 72, 179-194. 2012.

HWANG, S.; CHOI, S. Use of a lognormal distribution model for estimating soil water retention curves from particle-size distribution data. Journal of Hydrology, vol. 323, Issues 1–4, 2006.

IMMANUEL, C. D; DOYLE III, F. J; CORDEIRO, C. F; SUNDARAM, S. S. Population balance PSD model for emulsion polymerization with steric stabilizers. AIChE j. Vol.49:1392–404, 2003.

IMMANUEL, C.D; WANG, Y; BIANCO, N. Feedback controllability analysis and in-batch feedback control of particle size distribution in semi-batch emulsion polymerisation. Chem Eng Sci, 63(5): p.1205–1216, 2008.

IMMANUEL, C; DOYLE III, F. Computationally efficient solution of population balance models incorporating nucleation, growth and coagulation: application to emulsion polymerization. J. Chem. Eng. Sci. Vol.57, p.4415–4427, 2002.

JACOBS, R. A. Increased rates of convergence through learning rate adaptation, Neural Networks. Vol.1, p.295–307, 1988.

JIAO, B., LIAN, Z., GU, X. A dynamic inertia weight particle swarm optimization algorithm. Chaos Solitons & Fractals. 37, 698-705.2008.

JOHNSSON, J.; BLADH, H.; BENGTSSON, P. E. On the influence of bimodal size distributions in particle sizing using laser-induced incandescence. Applied Physics B: Lasers e Optics, vol.99, 2010.

KALOGIROU, S. A. Artificial intelligence for the modeling and control of combustion processes: a review. Progress in Energy and Combustion Science. v. 29, p. 515–566, 2003.

KATARE, S; KALOS, A; WEST, D. A Hybrid Swarm Optimizer for Efficient Parameter Estimation. IEEE, 2004.

KENNEDY, J. The particle swarm: Social adaptation of knowledge. Proc. IEEE International Conference on Evolutionary Computation. Indianapolis, 1997.

KENNEDY, J; EBERHART, R.C. New optimizer using particle swarm theory. International Symposium on Micro Machine and Human Science, 6. Proceedings. Nagoya, 1995.

KENNEDY, J; MENDES, R. Topological structure and particle swarm performance. In Proceedings of the Fourth Congress on Evolutionary Computation, 2002.

KIPARISSIDES, C. “Polymerization Reactor Modeling: A Review of Recent Developments and Future Directions”, Chemical Engineering Science, v. 51, pp. 1637-1659, 1996.

KIPARISSIDES, C; ACHILIAS, D. S; FRANTZIKINAKIS, C. E. The effect of oxygen on the kinetics and particle size distribution in vinyl chloride emulsion polymerization. Ind. Eng. Chem. Res. Vol.41, p.3097–3109, 2002.

KIRANYAZ, S.; INCE, T.; YILDIRM, A.; GABBOUJ, M. Evolutionary artificial neural networks by multi-dimensional particle swarm optimization. Neural Networks. Vol. 22, p.1448–1462, 2009.

KLEBER, C.; KOTZ, S. Statistical Size Distributions in Economics and Actuarial Sciences. Wiley series in probability and statistics, 2003.

KLINK, I. M.; PHILLIPS, R. J.; DUNGAN, S. R. Effect of emulsion drop-size distribution upon coalescence in simple shear flow: A population balance study. Journal of Colloid and Interface Science 353, 2011.

KURODA C.; KIM, J., Neural network modeling of temperature behavior in an exothermic polymerization process, Neurocomputing, v.43, p. 77 – 89, 2002.

LE VEQUE, R. J. Finite volume methods for hyperbolic problems. Cambridge: Cambridge University Press; 2002.

LENIN, K; MOHAN, M.R. Attractive and Repulsive Particle Swarm Optimization for Reactive Power Optimization. Journal of engineering and Applied Science. p.288-292. 2006.

LENZI, M. K., —Modelagem da Polimerização Simultânea de Estireno em Suspensão e Emulsão”. Dissertação de M.Sc., COPPE/UFRJ, Rio de Janeiro, RJ, Brasil, 2002.

LEUNG, S; TANG, Y; WONG, W. A hybrid particle swarm optimization and its application in neural networks. Expert Systems with Applications. Vol.39, p.395–405, 2012.

LIMPERT, E., STAHEL, W., ABBT, M. Lognormal distribution across the sciences: Keys and Clues. BioScience, vol.51, Issue 5, 2001.

LINDEN, R. — "Algoritmos Genéticos", 1ª Edição, Ed. Brasport, Brasil, 2006.

LIU, B; WANG, L; YI-HUI, J; TANG, F; HUANG, D. X. Improved particle swarm optimization combined with chaos. *Chaos Solitons & Fractals*, vol. 25, pp. 1261-1271, 2005.

MACHADO, F; LIMA, E. L; PINTO, J. C. Uma Revisão Sobre os Processos de Polimerização em Suspensão. *Polímeros: Ciência e Tecnologia*, vol. 17, nº 2, p. 166-179, 2007.

MAGGIORIS, D., "Use of CFD in Prediction of Particle Size Distribution in suspension Polymer Reactors". *Computer Chemical Engineering*, vol. 22, pp 315-322, 1998.

MAGGIORIS, D., GOULAS, A., ALEXOPOULOS, A.H., CHATZI, E.G., KIPARISSIDES, C., "Use of CFD in Prediction of Particle Size Distribution in Suspension Polymer Reactors", *Comp. Chem. Eng.*, vol. 22, Suppl., pp. S315-S322, 1998.

MAGURRAN A. E. *Ecological Diversity and its Measurement*. London: Croom Helm, 1988.

MALLIKARJUNAN, V; PUSHPAVANAM, S; IMMANUEL, C. D. Parameter estimation strategies in batch emulsion polymerization. *Chemical Engineering Science*, vol.65, pp.4967–4982. 2010.

MARINANGELO, G. — "Polimerização em emulsão de estireno e acrilato de butila com alto teor de sólidos: estudo experimental e modelagem matemática do processo em reator semicontínuo". Tese de doutorado, Escola Politécnica da Universidade de São Paulo, 2010.

MARING, H.; SAVOIE, D. L.; IZAGUIRRE, M. A.; CUSTALS, L.; REID, J. S. Mineral dust aerosol size distribution change during atmospheric transport, *J. Geophys. Res.*, 108, 2003.

MARIZ, I. de F. A.; LEIZA, J. R.; DE LA CAL, J. Competitive particle growth: A tool to control the particle size distribution for the synthesis of high solids content low viscosity latexes. *Chemical Engineering Journal*. Vol.168, p.938–946, 2011.

MARTIN, G. Consider soft sensors. *Chemical Engineering Progress*. July, 66, 1997.

MEISSNER, M; SCHUMUKER, M; SCHNEIDER, G. Optimized particle swarm optimization and its application to artificial neural network training. *BMC Bioinformatics*. Vol.7, 125. 2006.

MEYER, Paul. Probabilidade: Aplicações a Estatística. LTC, 2ª ed. 1983.

MEYER, T. & KEURENTJES, J. Handbook of Polymer Reaction Engineering. Ed. Wiley-VCH, 77p. 2005.

MIN, K. W; RAY, W. H. "On the Mathematical Modeling of Emulsion Polymerization Reactors", J. Macromol. Sci., p.177-255, 1974.

MINSKY, M.L; PAPERT, S.A. Perceptrons: Introduction to Computational Geometry. Expanded Edition the MIT Press, 1988 (1st edition: 1969).

MORBIDELLI, M., STORTI, G., CARRÁ, S. —Role of Micellar Equilibria on Modelling of Batch Emulsion Polymerization Reactors", Journal of Applied Polymer Sci., v.28, pp.901-919. 1983.

MORRISON, B. R; MAXWELL I. A., GILBERT R. G; NAPPER, D. H. Testing nucleation models for emulsion polymerization systems. ACS Symposium Series 492, p. 28-44: Polymer Latexes-Preparation, characterization and application. American Chemical Society, Washington DC, 1992.

MYERS, S. C.; MAJLUF, N. S. Corporate financing and investment decisions when firms have information that investors do not have. Journal of Financial Economics, vol. 13, Issue 2, 1984.

NASCIMENTO, C. A. O; GIUDICI, R; GUARDANI, R. Neural network based approach for optimization of industrial chemical processes. Computers & Chemical Engineering. Vol. 24, Issues 9–10, p. 2303–2314. 2000

NELE, M.; PINTO, J. C. Molecular-Weight Multimodality of Multiple Flory Distribution, Macromol. Theory Simul.11, 293-307, 2002.

NISHIURA, H. Early efforts in modeling the incubation period of infectious diseases with an acute course of illness. Emerging Themes in Epidemiology,4:2, 2007.

ODIAN, G. Principles of Polymerization, chapter 1, New Jersey, USA, John Wiley & Sons, 2004.

OLORUNDA, O; ENGELBRECHT, A. P. Measuring exploration/exploitation in particle swarms using swarm diversity. IEEE Congress on Evolutionary Computation.p.1128-1134, 2008.

OMRAN, M. G. H. D. Particle Swarm optimization for Pattern Recognition and Image processing. University of Pretoria, 2005.

OTTA, W. A Physical Explanation of the Lognormality of Pollutant Concentrations. Journal of the Air & Waste Management Association, vol. 40, Issue 10, 1990.

PAN, H; ZHENG, M. Particle Swarm-Simulated Annealing Fusion Algorithm and its Application in Function Optimization. International Conference on Computer and Software Engineering. p. 78-92. 2008.

PANT, M., RADHA, T., SINGH, V.P.: Particle Swarm Optimization: Experimenting the Distributions of Random Numbers. I: 3rd Indian Int. Conf. on Artificial Intelligence (IICAI 2007), India, pp. 412–420, 2007.

PARSOPOULOS, KE.; VRAHATIS M.N. Parameter selection and adaptation in Unified Particle Swarm Optimization. Mathematical and Computer Modeling. Jan, 2007.

PELEGA, M.; COLEB, M. B. Reinterpretation of Microbial Survival Curves. Critical Reviews in Food Science and Nutrition, vol. 38, Issue 5, 1998.

PINTO, J.C.; SANTOS, J.G.F.; WAY, D.V.; Melo Jr., P.A.; Nele, M. Analysis of Near Infrared Spectra During Methyl Methacrylate (MMA) Suspension Polymerizations. Macromolecular Symposia, v.299; issue 300, 57-65. 2011.

POLI, R; KENNEDY, J; BLACKWELL, T. Particle swarm optimization - An overview. Swarm Intelligence. Volume 1, Number 1, 33-57, 2007.

PRINDLE, J. “Dynamics and Stability of Emulsion Polymerization Reactor”, Phd Thesis - University of Wisconsin, 1989.

RADHA, M. P; SINGH, V. P. A Simple Diversity Guided Particle Swarm Optimization. IEEE Congress on Evolutionary Computation, 2007.

RALLO, R; FERRE-GINE, J; ARENAS, A; GIRALT, F. Neural virtual sensor for the inferential prediction of product quality from process variables. Computers and Chemical Engineering. Vol. 26 p.1735-1754, 2002.

RAMKRISHNA, D. Population Balance - Theory and Applications to Particulate Systems in Engineering. Academic Press, San Diego.2000.

RAUSAND, M.; HOYLAND, A. System Reliability Theory: Models, Statistical Methods, and Applications. John Wiley & Sons, 2004.

REIS, M. M.; ARAÚJO, P.H.H.; SAYER, C.; GIUDICI, R. Comparing Near and Raman Spectroscopy for on-line Monitoring of Emulsion Copolymerization Reactions. *Macromol. Symp*, 206, 165-178, 2004.

RICHARDS, M.; VENTURA, D. Dynamic sociometry in particle swarm optimization. *Proceedings of the Sixth International Conference on Computational Intelligence and Natural Computing*. North Carolina. pp. 1.557-1560, 2003

RIGOPOULOS, S. J. A Finite-element scheme for solution of the dynamic population balance equation. *AIChE J.* Vol.49:1127–39, 2003.

ROSS, S. M. Introduction to probability models. 7 ed. Academic Press. P. 693. 2000.

RUITER, C. J.; OESEBURG, F. A Comparison of Parameters Estimation Procedures for Experimental Bimodal Lognormal Particle Size Distributions and Grouped Data. *J. Aerosol Sci.*, Vol. 18, No. 4, pp. 431-444, 1987.

RUMELHART, D.E. & MCCLELLAND, J.L. —Parallel Distributed Processing: Explorations in the Microstructure of Cognition”, vols. 1 & 2, The MIT Press, 1986.

RYAN, T. P. Estatística Moderna para Engenharia. Tradução: Luiz Claudio de Queiros Faria. Elsevier. Rio de Janeiro, 2009.

SALDÍVAR, E; DAFINIOTIS, P; RAY, H. *Journal of Macromolecular Science- Reviews in Macromolecular Chemistry and Physics*, 27, 1998.

SARTWELL, P. E. The distribution of incubation periods of infectious disease. *American Journal Epidemiology*. 51 (3), 1950.

SCHWAAB, M., Avaliação de Algoritmos Heurísticos de Otimização em Problemas de estimativa de Parâmetros, Dissertação de Mestrado, COPPE/UFRJ. 2005

SEINFELD, J. H.; PANDIS, S. N. *Atmospheric Chemistry and Physics*, 1326 pp., John Wiley e Sons, New York, 1997.

SEXTON, R; DORSEY, R; JOHNSON, J. Toward a global optimum for neural networks: A comparison of the genetic algorithm and back propagation, *Decision Support Systems*. Vol. 22, Issue 2, p.171-186. 1998.

SHI, Y.; EBERHART R. A modified particle swarm optimizer. Proc. IEEE Int. Conf on Evolutionary Computation, p.69-73, 1998.

SHI, Y.; EBERHART R. Empirical Study of Particle Swarm Optimization. IEEE, 1999.

SINGER, D. A. The lognormal distribution of metal resources in mineral deposits. Ore Geology Reviews, vol. 55, 2013.

SINGH, V.; GUPTA, I.; GUPTA, H. O. ANN-based estimator for distillation using Levenberg Marquardt approach. Engineering Applications of Artificial Intelligence. Vol. 20, p. 249-259. 2007.

SOOD, A. Particle Size Distribution Control in Emulsion Polymerization. Journal of Applied Polymer Science, Vol. 92, 2004.

STEPNIEWSKI, S. W; KEANE, A. J. Pruning back-propagation neural networks using modern stochastic optimization techniques. Neural Computer Applied, vol. 5, no. 2, pp. 76-98, 1997.

SU, C. L; YANG, S. M; HUANG, W. L. A two-stage algorithm integrating genetic algorithm and modified Newton method for neural network training in engineering systems. Expert Systems with Applications 38: 2011.

SWEETMAN, S. J; IMMANUEL, A. C. D; MALIK, A. T. I; EMMETT, B. S; WILLIAMS, C. N. 18th European Symposium on Computer Aided Process Engineering, 2008.

TAYLOR, J. A.; JAKEMAN, A. J.; SIMPSON, R. W. Modeling distributions of air pollutant concentrations—I. Identification of statistical models. Atmospheric Environment, vol. 20, Issue 9, 1986.

THANGARAJ, R; PAMT, M. ABRAHAM, A. BOUVRY, P. Particle swarm optimization: Hybridization perspectives and experimental illustrations. Applied Mathematics and Computation. Vol.217, p.5208-5226, 2011.

THICKETT, S. C. & GILBERT, R. G. Emulsion polymerization: State of the art in kinetics and mechanisms. Polymer. v.48, 6965-6991, 2007.

THOMAS, N; REED, M. A hybrid algorithm for continuous optimization. IEEE Congress on Evolutionary Computation. 2009.

THOMSON, R. C; YASZEMSKI, M. J; POWERS, J. M. Fabrication of biodegradable polymer scaffolds to engineer trabecular bone. J. Biomed. Mater. Res. 44, p. 314-321. 1995.

TIAN, Y.; ZHANG, J.; MORRIS, J., Optimal control of a batch emulsion copolymerisation reactor based on recurrent neural networks models, *Chemical Engineering and Processing*, v. 41, p. 531-538, 2002.

TIAN, K.; LIU, F. S.; THOMSON, K. A.; SNELLING, D.R.; SMALLWOOD, G.J.; WANG, D.S. Distribution of the number of primary particles of soot aggregates in a non premixed laminar flame. *Combustion Flame*, vol.138, issue 1-2, 2004.

TOBIAS, Melvin. Using the Log-normal Distribution in Analyzing Aerosols: the Mathematical Reasoning Underlying the Various Diameters Used and the Plotting Procedure. OAK RIDGE NATIONAL, LABORATORY, Tennessee. 1993.

TRELEA I. C. The particle swarm optimization algorithm: convergence analysis and parameter selection. *Information Processing Letters*. Vol.85, p.317-325, 2003.

TRIOLA, Mario F. Introdução à estatística. LTC. Rio de Janeiro. LTC. 10 ed. 696 p. 2008.

TSEN, A. Y-D; JANG, S. S; WONG, D. S. H; JOSEPH, B. Predictive control of quality in batch polymerization using hybrid ANN models. *AIChE Journal*. Vol.42, p.455–465, 1996.

TSI – TRUST, SCIENCE, INNOVATION. Aerosol Statistics Lognormal Distribution and $dN/d\log D_p$. Disponível em: http://www.tsi.com/uploadedFiles/_Site_Root/Products/Literature/Application_Notes/PR-001-RevA_Aerosol-Statistics-AppNote.pdf. Acessado em 10/08/2013.

TULI, A.; KOSUGI, K.; HOPMANS, J.W. Simultaneous scaling of soil water retention and unsaturated hydraulic conductivity functions assuming lognormal pore-size distribution. *Advances in Water Resources*, vol. 24, Issue 6, 2001.

ULIANA, M. Soluções de equação de balanço populacional pelo método de classes com aplicação a processo de polimerização em suspensão. Dissertação de mestrado apresentada ao Programa de Pós-Graduação em Engenharia Química da Escola Politécnica da Universidade de São Paulo, 2007.

VALE, H.M., MCKENNA, T.F. Solution of population balance equations for emulsion polymerization: Zero–One and Zero–One–Two Systems. *Ind. Eng. Chem. Res.* 46, 643–654. 2007.

- VALE, H.M; McKENNA, T. F. Modeling particle size distribution in emulsion polymerization reactors. Progress in polymer Science. 30, 1019-1048, 2005.
- VAN DEN BERG, F. An Analysis of Particle Swarm Optimizers. PhD thesis, Department of Computer Science, University of Pretoria, Pretoria, South Africa, 2002.
- VAN DEN BERGH, F. Particle swarm weight initialization in multi-layer perceptron artificial neural networks. Development and Practice of Artificial Intelligence Techniques pp. 41-45. Durban, South Africa, 1999.
- VAN DEN BERGH, F. e ENGELBRECHT, A.P. A study of particle swarm optimization particle trajectories. Information Science. Vol.176, p.937-971, 2006.
- VAN OUYEN, A.; NIENHUIS, B. Improving the convergence of the back-propagation algorithm, Neural Network. Vol. 5. Issue 4, p.465–471, 1992.
- VON ZUBEN, F; CASTRO, L. N. Redes Neurais Artificiais. DCA/FEEC/Unicamp. Disponível em: ftp://ftp.dca.fee.unicamp.br/pub/docs/vonzuben/ia006_03/topico5_03.pdf, acessado em 20/08/2010.
- ZEAITER, J; ROMAGNOLI, J. A; GOMES, V. G. Online control of molar mass and particle-size distributions in emulsion polymerization. AIChE Journal. Volume 52, Issue 5, pages 1770–1779, 2006.
- WANG, D; LU, W. Forecasting of ozone level in time series using MLP model with a novel hybrid-training algorithm. Atmospheric Environment. Vol. 40, p.913–924, 2006.
- WANG, S. The unpredictability of standard back propagation neural networks in classification applications, Management Science. v.41, p. 555-559, 1995.
- WARSON, H; FINCH, C. A. Applications of Synthetic Resin Latices, 3 vols. JohnWiley, Chichester, UK, 2001.
- WEBER, S. H. Desenvolvimento de uma Função Densidade de Probabilidade Multimodal Aplicável à Ciência Florestal. Tese de Doutorado apresentada ao Programa de Pós-Graduação em Engenharia Florestal da Universidade Federal do Paraná. Curitiba - Brasil, 2011.
- WEIRS, M. K. A method for self-determination of adaptive learning rates in back propagation, Neural Networks. Vol.4, p. 371–379, 1991.

WHITE, T.; PAGUREK, B.; Towards Multi-Swarm Problem Solving in networks. Proceedings Third International Conference on Multi-Agent Systems (ICMAS '98), p. 333-340, 1998.

WILAMOWSKI, B. M.; IPLIKCI, S.; KAYNAK, O.; EFE, M. O. An algorithm for fast convergence in training neural networks, in: Proceeding of the International Joint Conference on Neural Network. Vol3, p.1778-1782, 2001.

YAO, X. "Evolving artificial neural networks", Proceedings of the IEEE, vol. 87, pp. 1423-1447, 1999.

YU, J; WANG, S; XI, L. Evolving artificial neural networks using an improved PSO and DPSO. Neurocomputing. Vol. 71, Issues 4–6, p.1054–1060, 2008.

YUAN, Y.; YI, H.; SHUAI, Y.; LIU, B.; TAN, H. Inverse problem for aerosol particle size distribution using SPSO associated with multi-lognormal distribution model. Atmospheric Environment, vol. 45, 2011.

ZANCHETTIN, C.; LUDERMIR, T. Global Optimization Methods for Designing and Training Feed forward Artificial Neural Networks. Advances in Neural Networks, Vol. 14 [S1], p.328-337, 2007.

ZENDER, Charlie. Particle Size Distributions: Theory and Application to Aerosols, Clouds, and Soils. Online: <http://dust.ess.uci.edu/facts>, Updated: Mon 30th Aug, 2010.

ZHAN, Z-H.; ZHANG, J; LI, Y; CHUNG, H.S-H. Adaptive particle swarm optimization. IEEE Transactions on Systems Man, and Cybernetics — Part B: Cybernetics, Vol.39 (6). p. 1362-1381, 2009.

ZHANG, C.; SHAO, H.; LI, Y. Particle swarm optimization for evolving artificial neural network. In: Systems, Man, and Cybernetics, 2000 IEEE International Conference on. Vol. 04, p. 3562-3563, 2000.

ZHANG, J. Batch-to-batch optimal control of a batch polymerization process based on stacked neural network models. Chemical Engineering Science. Vol.63, p.1273-1281, 2008.

ZHANG, J. R; ZHANG, J; LOK, T. M; LYU, M. R. A hybrid particle swarm optimization–back-propagation algorithm for feed forward neural network training. Applied Mathematics and Computation. Vol. 185, p.1026–1037, 2007.

ZHANG, J; MARTIN, E. B; MORRIS, A. J; KIPARISSIDES, C. Inferential Estimation of Polymer Using Stacked Neural Networks. *Computer and Chemical Engineering*.1998, 21. 1025-1030, 1998.

ZHANG, J; MORRIS, A. J; MARTIN, E. B; KIPARISSIDES, C. Estimation of Impurity and Fouling in Batch Polymerization Reactors Through the Application of Neural Networks. *Computers and Chemical Engineering*, 23, 301-314, 1999.

ZHANG, M., RAY, W., H. —Modeling of Living Free-Radical Polymerization Processes. I. Batch, Semi batch, and Continuous Tank Reactors”, *Journal of Applied Polymer Science*, Vol. 86, p. 1630-1662, 2002.

ZHANG, Z.; FRIEDRICH, K. Artificial neural networks applied to polymer composites: a review. *Composites Science and Technology*. Vol.63, p.2029–2044, 2003.

ZHAO, S. J.; XU, Y. M. Levenberg Marquardt algorithm for nonlinear principal component analysis neural network trough inputs training, in: *Fifth World Congress on Intelligent Control and Automations*, Vol.4, p. 3278-3281, 2004.

ZHENG, Y. G; ZHANG, H. W; CHEN, Z. Combined Stochastic Diffusion and Mean-Field Model for Grain Growth. *Computational Mechanics*, 2009.

ZUBITUR, M.; ASUA, J. M.; Factors affecting kinetics and coagulum formation during the emulsion copolymerization of styrene/butyl acrylate. *Polymer*.42, p.5979-5985, 2001.

APÊNDICE A – Modelagem da Distribuição de Tamanho de Partícula fazendo uso das Equações de Balanço Populacional

A Equação (26) representa o balanço populacional típico para o processo de polimerização em emulsão (IMMANUEL *et. al* 2003), onde r representa o raio das partículas e $F(r,t)$ é a função densidade populacional das partículas, definida de modo que $F(r,t)dr$ equivale à quantidade (mols) de partículas com raio entre r e $r + dr$:

$$\frac{\partial F(r,t)}{\partial t} + \frac{\partial}{\partial t} \left(F(r,t) \frac{dr}{dt} \right) = \mathcal{R}_{nuc} \delta(r - r_{nuc}) + \mathcal{R}_{coal} \quad (26)$$

Nesta equação, o termo dr/dt representa a taxa de crescimento das partículas de raio r , o termo \mathcal{R}_{nuc} representa a taxa de formação de novas partículas de raio igual a r_{nuc} (a restrição do tamanho das novas partículas é representada pela função $\delta(r - r_{nuc}) = 0$ se $r \neq r_{nuc}$ e $\delta(r - r_{nuc}) = 1$ se $r = r_{nuc}$) e \mathcal{R}_{coal} representa a taxa de coalescência das partículas.

A taxa de coalescência das partículas pode ser separada em dois termos: um para a formação e outro para o consumo das partículas devido à coalescência $\mathcal{R}_{coal} = \mathcal{R}_{form} + \mathcal{R}_{cons}$.

A taxa de formação de partículas de tamanho r devido à coalescência pode ser dada pela Equação (27) (IMMANUEL *et. al* 2003). Nesta equação, $r/(2^{1/3})$ limita o diâmetro máximo de partículas envolvidas na formação de uma partícula de tamanho r , através da coalescência:

$$\mathcal{R}_{form}(r,t) = \frac{1}{V_{aq}} \int_{r'=r_{nuc}}^{\frac{r}{2^{1/3}}} B(r',r'') F(r',t) F(r'',t) \frac{r^2}{(r^3 - (r')^3)^{2/3}} dr' \quad (27)$$

A taxa de consumo de partículas devido à coalescência é dada pela Equação (28):

$$\mathcal{R}_{cons}(r,t) = \frac{1}{V_{aq}} \int_{r_{nuc}}^{r_{máx}} B(r,r') F(r,t) F(r',t) dr' \quad (28)$$

onde $B(r,r')$ é o coeficiente da taxa de coalescência de duas partículas de raios r e r' e V_{aq} é o volume da fase aquosa.

Da mesma forma, a taxa de nucleação de partículas pode ser separada em dois termos, um para a nucleação homogênea e outro para a nucleação micelar $\mathcal{R}_{nuc} = \mathcal{R}_{micelar} + \mathcal{R}_{homogênea}$. Considerando a hipótese de que os radicais apenas podem entrar em uma micela para nuclear uma nova partícula após atingir um tamanho mínimo z , a taxa de nucleação micelar pode ser dada pela Equação (29).

$$\mathcal{R}_{micelar} = \sum_{m=z}^{j_{crit}-1} \sum_{i=A,B} k_{emic,i} P_{wi} [R_m^*]_{aq} N_{mic} \quad (29)$$

onde $k_{emic,i}$ é o coeficiente da taxa de captura de radicais pelas micelas, $[R_m^*]_{aq}$ é a concentração de radicais na fase aquosa de tamanho m entre z e $j_{crit}-1$ e P_{wi} é a probabilidade de que o radical na fase aquosa seja terminado por um mero do tipo i ;

Considerando a hipótese de que os radicais, ao atingirem o tamanho j_{crit} precipitam formando uma nova partícula, a taxa de formação de partículas por nucleação homogênea é obtida por:

$$\mathcal{R}_{homogenea} = k_{p,av}^{aq} \left[R_{j_{crit}-1}^* \right]_{aq} V_{aq} \quad (30)$$

onde $k_{p,av}^{aq}$ é o coeficiente da taxa de propagação na fase aquosa para pseudo-homopolimerização e $[R_{j_{crit}-1}^*]_{aq}$ é a concentração de radicais de tamanho $j_{crit}-1$ na fase aquosa.

Contudo, um balanço populacional completo em um processo de polimerização em emulsão deve considerar o número de radicais ativos existentes em cada partícula n , que assume valores discretos. Assim, a função densidade de partículas, para um processo de copolimerização de dois monômeros assumiria a forma $F_{n_A, n_B}(r, t)$, onde n_A e n_B , os números de radicais tipo A e tipo B presentes na partícula, podem assumir valores discretos $(0, 1, 2, \dots)$, resultando em um sistema de equações de balanço para as diferentes combinações de n_A e n_B . De acordo com Marinangelo (2010), o modelo pode ser consideravelmente simplificado utilizando-se um número médio de radicais por partícula para cada faixa de tamanho. Em processos de copolimerização pode-se empregar ainda a hipótese de pseudo-homopolimerização, resultando em um único número médio de radicais por partícula em função do tamanho das partículas, $\bar{n}(r, t)$. Assim, o número médio de radicais por partícula pode ser obtido também através de uma equação de balanço populacional, como mostra a Equação (31):

$$\frac{\partial(\bar{n}(r,t)F(r,t))}{\partial t} + \frac{\partial}{\partial r}\left(\bar{n}(r,t)F(r,t)\frac{dr}{dt}\right) = \mathfrak{R}_{ent}(r) + \mathfrak{R}_{des}(r) + \mathfrak{R}_{ter}(r) + \mathfrak{R}_{coag}(r) \quad (31)$$

Nesta equação, os termos do lado direito representam as contribuições da entrada, da saída, da terminação dos radicais e da coagulação das partículas. O segundo termo da esquerda representa o efeito do crescimento da partícula. Se negligenciadas as contribuições devido ao crescimento das partículas e coalescência, o balanço populacional para os radicais pode ser reduzido a um conjunto de equações diferenciais ordinárias (IMMANUEL *et. al* 2003). Neste caso, é possível utilizar como alternativa uma única partícula de raio r , como volume de controle para o cálculo do número médio de radicais por partícula de mesmo raio. Obtendo, assim, a Equação (32) para o balanço de radicais nesta partícula:

$$\frac{\partial(\bar{n}(r,t))}{\partial t} = \mathfrak{R}'_{ent}(r) + \mathfrak{R}'_{des}(r) - 2\mathfrak{R}'_{ter}(r) \quad (32)$$

As taxas de entrada, saída e terminação dos radicais para uma partícula de tamanho r são dadas pelas Equações (33), (34) e (35), respectivamente:

$$\mathfrak{R}'_{ent}(r) = \sum_{m=z}^{j_{crit}-1} \sum_i^{A,B} k_{e,i}^m(r) \cdot p_{wi} \cdot [R_m^*]_{aq} \quad (33)$$

$$\mathfrak{R}'_{des}(r) = \sum_i^{A,B} k_{dMj}(r) \cdot \frac{\sum_i^{A,B} k_{trij} \cdot p_{wi} \cdot \bar{n}(r,t) [M_j]_p}{k_{dMj}(r) + \sum_l^{A,B} k_{pjl} \cdot [M_j]_p} \quad (34)$$

$$k_{dMj} = 3.D_{wj} / r^2 \quad (35)$$

onde $k_{e,i}^m$ é o coeficiente da taxa de captura de radicais de tipo i e tamanho m pelas partículas e k_{dMj} é o coeficiente da taxa de saída de radicais, dado por:

$$\mathfrak{R}'_{ter}(r) = \frac{k_{tav} \cdot \bar{n}(r,t)^2}{V_p(r) \cdot N_A} \quad (36)$$

com D_{wj} sendo o coeficiente de difusão do monômero j na fase aquosa, k_{pjl} é o coeficiente da taxa de propagação de um radical terminado em j com monômero do tipo l , $V_p(r)$ é o volume de uma partícula de tamanho r , k_{trij} é o coeficiente da taxa de transferência de cadeia de um radical terminado em i para um monômero do tipo j e k_{tav} é o coeficiente da taxa de terminação para pseudo-homopolimerização, dado por:

$$k_{tav} = \sum_i^{A,B} \sum_j^{A,B} p_i \cdot p_j \cdot k_{tij} \quad (37)$$

Finalmente, sendo r o raio das partículas não inchadas de monômero, a taxa de crescimento das partículas, para o caso de processo de copolimerização e dadas às simplificações anteriores, pode ser dada pela Equação (38):

$$\frac{dr}{dt} = \frac{1}{4\pi r^2 \rho_{pol}} \sum_i^{A,B} \sum_j^{A,B} k_{pij} p_i \frac{\bar{n}(r,t)}{N_A} [M_j]_p MM_j \quad (38)$$

onde $[M_j]_p$ é a concentração do monômero j nas partículas de látex, k_{pij} é o coeficiente da taxa de propagação de um radical terminado em i com monômero do tipo j , N_A é o número de Avogadro, MM_j é a massa molecular do monômero j , ρ_{pol} é a densidade do polímero e p_i é a probabilidade que um radical nas partículas seja de tipo i .

APÊNDICE B – Efeitos das condições de operação sobre a conversão e diâmetro médio de partícula

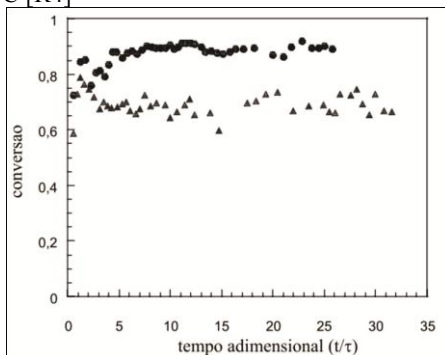
Como comentado desde o princípio deste trabalho, os dados experimentais utilizados na validação da estratégia proposta, são oriundos de Araújo (1999). Neste trabalho, o autor investiga o efeito das condições de operação sobre a distribuição do tamanho de partícula na copolimerização em emulsão de acetato de vinila e Veova10 em um reator loop contínuo com alto teor em sólidos.

Uma síntese dos efeitos das condições de operação sobre a conversão e diâmetro médio de partícula é apresentada a seguir, de acordo com as conclusões de Araújo (1999):

Tempo de residência - O autor observa que não há grandes variações da conversão e do diâmetro médio de partícula quando alterado o tempo de residência de 13,3 mim para 26,5 mim.

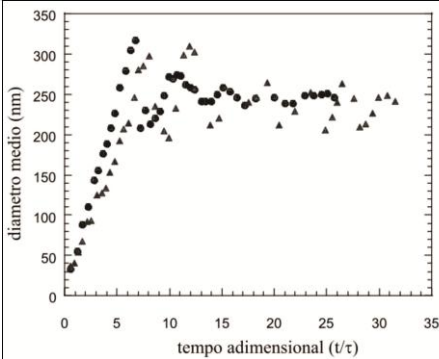
Temperatura - Verificou-se um forte declínio na conversão quando se diminui a temperatura da reação de 60°C para 50°C. Além disso, na temperatura de 50°C a taxa inicial de crescimento médio das partículas é menor, sendo que nessa condição o diâmetro médio permaneceu oscilante devido as renucleações. Estes resultados referentes à influência da temperatura sobre a conversão e diâmetro médio de partícula podem ser observados nas Figuras (29) e (30), respectivamente.

Figura 29 Efeito da temperatura na conversão (●) $T = 60^{\circ}\text{C}$ [R2], (▲) $T = 50^{\circ}\text{C}$ [R4]



Fonte: Araújo (1999)

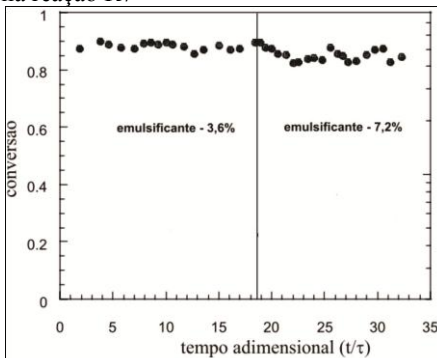
Figura 30 Efeito da temperatura no diâmetro médio de partícula (●) $T = 60^\circ \text{C}$ [R2], (▲) $T = 50^\circ \text{C}$ [R4]



Fonte: Araújo (1999)

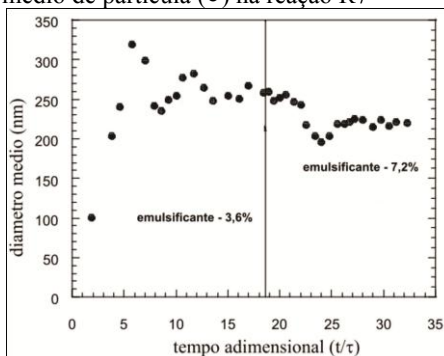
Concentração de emulsificante – Como ilustra a Figura (31) não houve variação significativa da conversão quando alterada a concentração de emulsificante. Já o diâmetro médio das partículas diminui quando a concentração de emulsificante aumenta em função do aumento da taxa de nucleação, como pode ser observado da Figura (32).

Figura 31 Efeito da concentração de emulsificante na evolução da conversão (●) na reação R7



Fonte: Araújo (1999).

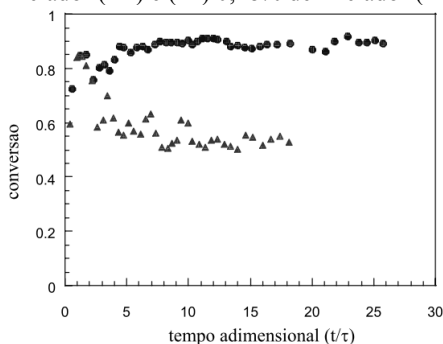
Figura 32 Efeito da concentração de emulsificante na evolução do diâmetro médio de partícula (●) na reação R7



Fonte: Araújo (1999).

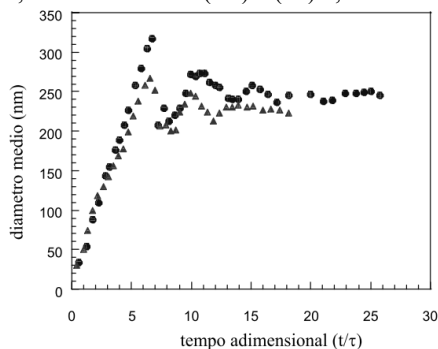
Concentração de iniciador – A redução na concentração de iniciador provocou uma redução da conversão da mesma ordem de grandeza (Figura 33). Verifica-se também uma diminuição do diâmetro médio de partícula quando a concentração de iniciador é menor, (Figura 34), visto que essa menor concentração de iniciador induz a formação de partículas menores.

Figura 33 Efeito da concentração de iniciador sobre a conversão. (●) 0,30% de iniciador (R2) e (▲) 0,18% de iniciador (R8)



Fonte: Araújo (1999)

Figura 34 Efeito da concentração de iniciador sobre o diâmetro médio. (●) 0,30% de iniciador (R2) e (▲) 0,18% de iniciador (R8)



Fonte: Araújo (1999).

Estas considerações são relevantes, já que explicitam a relação que existe entre variáveis de entrada e saída da rede neural ajudando na escolha dos dados das reações que farão parte do treinamento e validação da rede e na interpretação dos resultados.

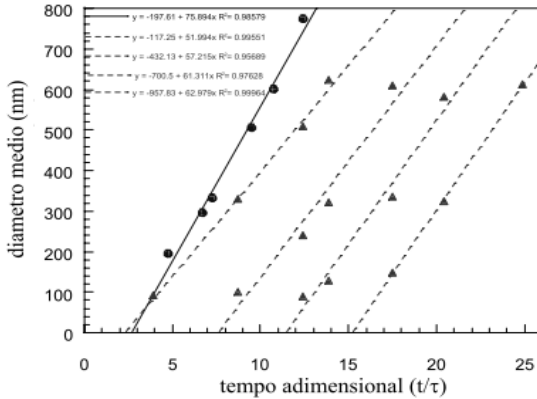
APÊNDICE C – Interpolações lineares da curvas de distribuições de tamanho de partícula

A distribuição de tamanho de partícula em látex polimérico pode ser determinada por vários métodos. Os métodos utilizados normalmente em medidas —offline— são os métodos de microscopia (incluindo ótica, por transmissão, varredura, transmissão de varredura, e tunelamento) e os de fracionamento por sedimentação (por gravidade e centrifugação). Estes métodos não são adequados para medidas em linha por requererem um preparo prévio da amostra, ou necessitarem de muito tempo para obtenção das medidas (ARAÚJO, 1999).

Por serem as DTP das amostras das reações realizadas muito polidispersas, o tempo necessário para analisar cada amostra na centrífuga disco era de aproximadamente 3 horas. Como cada reação resultava em mais de 40 amostras, seria impossível analisar todas as amostras por centrífuga disco. Por conta disso, Araújo, (1999) optou por determinar a DTP em $\pm 25\%$ dessas amostras, realizando em paralelo a determinação do diâmetro médio de partículas por Coulter em todas as amostras. Este procedimento permitiu ao autor verificar a concordância existente entre a evolução do diâmetro médio e renucleações que eram verificadas quando da determinação das distribuições de tamanho de partícula. De forma mais elucidativa, segundo o autor, verificou-se que, quando o diâmetro médio das partículas começava a diminuir era porque estava iniciando uma nova renucleação e quando o diâmetro aumentava era porque a renucleação havia cessado.

Outra constatação importante relatada por Araújo (1999) é a de que o diâmetro das partículas cresce linearmente, como pode ser verificado na Figura (35). De acordo com Burden e Faires (2008) esta informação permite que seja realizada a interpolação linear das DTPs nos casos em que não foram realizadas determinações experimentais. Os resultados das interpolações lineares realizadas são discutidos a seguir.

Figura 35 Velocidade de crescimento das diferentes populações de partícula em (▲) T = 60oC (R2) e (●) T = 50oC (R4)



Fonte: (ARAÚJO, 1999)

Observando os dados experimentais apresentados por Araújo, (1999), verificou-se que são breves os intervalos de amostragens em que não foram realizadas as determinações experimentais da DTPs, logo as interpolações podem ser realizadas sem perda considerável das características das curvas de distribuição de cada amostra. Para cada um dos pares de pontos $(d_i, f(d_i))$, das DTPs estabelecidos experimentalmente realizou-se uma interpolação linear, gerando um polinômio de grau um, definido pela Equação (39)

$$P(d_i) = h_0 + h_1 \cdot d_i \quad (39)$$

Estes polinômios permitiram a determinação dos pontos sobre a curva de DTP nas amostras ignoradas experimentalmente. Vale ressaltar ainda que as amostragens foram realizadas em intervalos de tempo semelhantes, logo é possível admitir que as curvas sejam igualmente espaçadas, simplificando a determinação dos diâmetros d_i de cada distribuição das n amostras interpoladas.

A interpolação é uma técnica que tem com objetivo avaliar quantitativamente valores funcionais de uma função f , sendo conhecido apenas um pequeno número de valores funcionais em argumentos denominados pontos bases, na forma $(x_i, f(x_i))$, $0 \leq i \leq n$, isto é, deseja-se construir uma função que forneça uma estimativa ao valor de $f(x)$, sendo $x \neq x_i$, com $0 \leq i \leq n$.

Os polinômios são as funções interpoladoras mais populares não só por suas propriedades algébricas, mas, sobretudo pela justificativa

fornecida pelo teorema de Weirstrass, que de fato garante a existência de um polinômio capaz de aproximar a função f tão bem quanto se queira (BURDEN e FAIRES, 2008).

No caso particular das interpolações requeridas nesse trabalho deve-se considerar que o polinômio interpolador admitido terá grau um, visto que são dados somente dois pontos distintos de uma função $f(d) : (d_a, f(d_a)) e (d_b, f(d_b))$, assim:

$$P(d) = h_0 + h_1 \cdot d \quad (40)$$

Para determinar os coeficientes h_0 e h_1 deve-se admitir que:

$$\begin{cases} P(d_a) = f(d_a) \\ P(d_b) = f(d_b) \end{cases}$$

ou seja, basta resolver o seguinte sistema linear:

$$\begin{cases} h_1 \cdot d_a + h_0 = f(d_a) \\ h_1 \cdot d_b + h_0 = f(d_b) \end{cases} \quad (41)$$

O determinante da matriz dos coeficientes é diferente de zero, sempre que $d_a \neq d_b$, logo para pontos distintos o sistema tem solução única.

Sendo assim, dados dois pontos distintos de uma função $f(d)$, deseja-se calcular o valor de f para um determinado valor de d entre d_a e d_b , sendo a e b pontos pertencentes as curvas determinadas experimentalmente.

Considerando que as amostragens foram realizadas em tempos semelhantes, admite-se que as curvas das k distribuições interpoladas estejam igualmente espaçadas. Assim, os diâmetros $d(n)$, com $n = 1 \dots k$ podem ser definidos pela Equação (42).

$$d(n) = \left| \frac{d_b - d_a}{k + 1} \right| + d(n - 1) \quad (42)$$

Por fim, como existem i pontos sobre cada curva que representa uma distribuição, com $i = 1 \dots n_{pt}$, a Equação (42) pode ser reescrita pela Equação (43):

$$d_n(i) = \left| \frac{d_b(i) - d_a(i)}{k + 1} \right| + d_{n-i}(i) \quad (43)$$

No intuito de demonstrar a adequação da proposta de interpolação linear com base no pressuposto do crescimento linear do diâmetro das partículas, foram realizados vários testes com diferentes distribuições,

sendo estas unimodais, bimodais e trimodais. A Tabela (16) apresenta pontos sobre a curva de DTP das amostras 15 e 19 na reação R2, determinados experimentalmente. A partir desses pares de pontos foram realizadas as interpolações possibilitando assim a determinação dos pontos pertencentes às distribuições das amostras 16, 17 e 18 na reação R2.

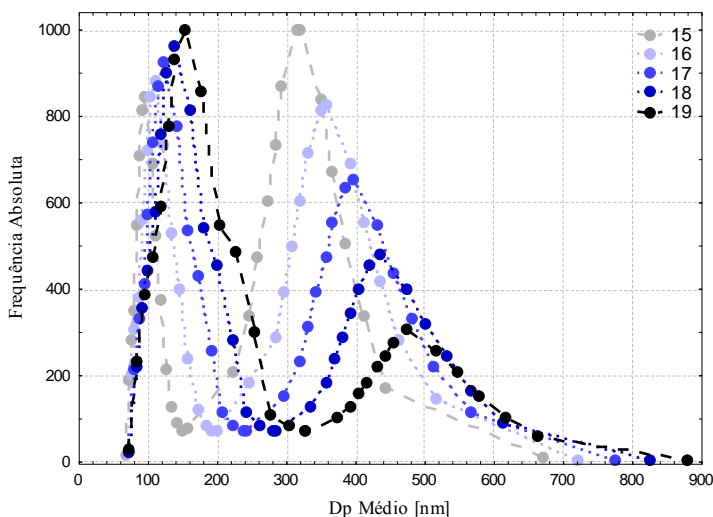
Tabela 16 Dados da distribuição de tamanho de partícula das amostras 15 e 19 na reação R2

I	Amostra 15		Amostra 19	
	Dp (nm)	Freq. (d)	Dp (nm)	Freq. (d)
1	71	17	75	25
2	75	189	86	230
3	78	280	97	383
4	82	350	108	471
5	86	544	119	592
6	90	706	130	772
7	94	811	141	927
8	97	843	154	998
9	107	690	179	854
10	114	518	204	544
11	122	371	229	483
12	129	212	254	299
13	137	126	279	106
14	144	86	304	80
15	151	69	327	70
16	158	73	327	70
17	224	205	373	99
18	248	337	396	128
19	260	469	407	157
20	273	600	418	181
21	284	733	432	215
22	295	865	443	244
23	316	995	455	273
24	320	1000	477	303
25	352	834	517	253
26	366	668	549	203
27	388	502	580	153
28	412	336	618	103
29	446	170	666	57
30	672	5	880	4

Fonte: Desenvolvido pelo autor

O resultado dessas interpolações pode ser verificado na Figura (36) que apresenta as curvas de distribuição de diâmetro de partícula nas amostras 16, 17 e 18. Em síntese, os pontos que geram as curvas de distribuições interpoladas estão em harmonia com o comportamento das curvas de distribuição das amostras 15 e 19, conferindo confiabilidade à técnica empregada.

Figura 36 Evolução das distribuições de tamanho de partícula entre as amostras 15 e 19 na reação R2



Fonte: Desenvolvido pelo autor

Este procedimento matemático permitiu que o conjunto de dados que descrevem a evolução das DTPs nas reações investigadas por Araújo (1999) pudesse ser triplicado, não podendo ser realizado somente em casos onde não houve muita concordância entre os pontos que compõem distribuições subsequentes oriundas de determinações experimentais, como também em casos que exigissem extrapolações já que não foram realizadas determinações experimentais de DTPs no início das reações, anteriores a dez tempos de residência.

O aumento considerável no conjunto de dados referentes às distribuições é fundamental para o sucesso da estratégia de inferência de DTPs por RNs. Isso porque, caso a quantidade de dados seja pequena, a rede não conseguirá criar um modelo suficientemente representativo

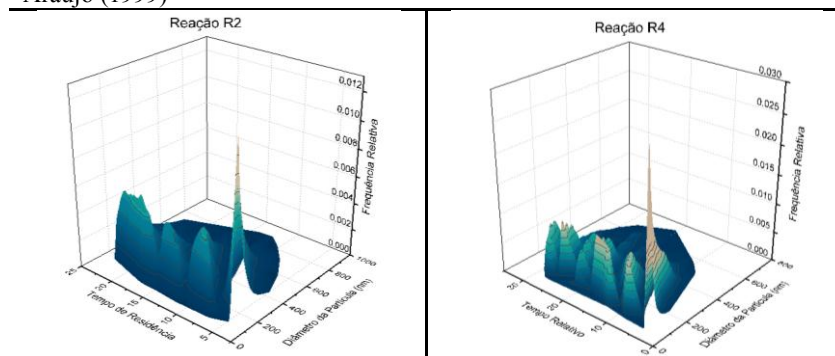
para se ter um bom desempenho quando aplicado em situações reais após o seu desenvolvimento. Além disto, os dados devem englobar todos os aspectos do problema em questão, a fim de que o modelo criado seja genérico (HAYKIN, 2001).

APÊNDICE D – Comportamento da distribuição de tamanho de partícula ao longo das reações

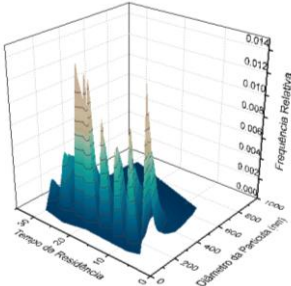
O sucesso do procedimento matemático de interpolação linear, discutida no Apêndice A, permitiu que fossem geradas superfícies tridimensionais [Tempo de Residência X Diâmetro da Partícula X Frequência Relativa] que ilustram o comportamento das DTPs no decorrer das reações.

A Figura (37) apresenta as superfícies geradas pelas DTPs para todas as reações realizadas por Araújo (1999) e que foram utilizadas neste trabalho. Uma importante constatação, decorrente da observação das superfícies geradas, é a clareza dos picos e vales gerados a cada renucleação, como pode ser observado na reação R8. Além disso, a nucleação inicial produz uma moda que se sobre sai as demais, o que era esperado pelo excesso de monômero oriundo da pré-alimentação do reator. A fim de reafirmar a escolha da reação R5 como validação da Rede Neural implementada, percebe-se que a referida reação apresenta muitas perturbações, não podendo ser classificada como uma reação simples de ser predita.

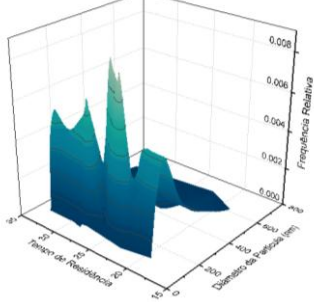
Figura 37 Superfícies geradas a partir das DTPs, para as reações realizadas por Araújo (1999)



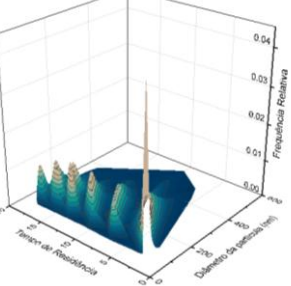
Reação R5



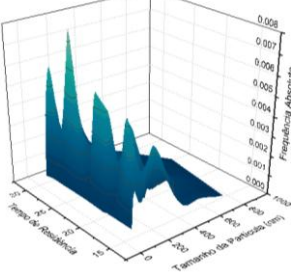
Reação R7



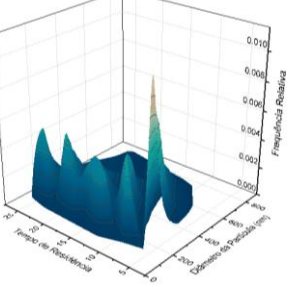
Reação R8



Reação R9



Reação R10



Fonte: Desenvolvido pelo autor

APÊNDICE E – Matriz de correlação entre parâmetros ajustados da Função Lognormal Bimodal

As Tabelas (17) e (18) apresentam a matriz de correlação (coeficiente de correlação de Pearson) entre os parâmetros ajustados da Função Lognormal Bimodal, para as reações R5 e R8, respectivamente. Considerando um nível de significância de 5%, os pontos destacados indicam correlação entre os parâmetros.

Tabela 17 Matriz de Correlação entre os parâmetros ajustados da Função Lognormal Trimodal [Reação – R5]

	M1	M2	M3	D1	D2	D3	P1	P2	P3
M1	1.00								
M2	0.40	1.00							
M3	0.21	0.55	1.00						
D1	-0.44	0.35	0.33	1.00					
D2	-0.23	-0.59	-0.03	-0.45	1.00				
D3	-0.04	-0.05	-0.53	-0.32	0.11	1.00			
P1	0.06	0.41	0.25	0.65	-0.67	-0.52	1.00		
P2	-0.05	-0.23	0.04	-0.61	0.69	0.35	-0.83	1.00	
P3	-0.01	-0.27	-0.47	-0.03	-0.08	0.25	-0.23	-0.35	1.00

Fonte: Desenvolvido pelo autor

Tabela 18 Matriz de Correlação entre os parâmetros ajustados da Função Lognormal Trimodal [Reação – R8]

	M1	M2	M3	D1	D2	D3	P1	P2	P3
M1	1.00								
M2	0.65	1.00							
M3	0.54	0.74	1.00						
D1	0.45	0.63	0.21	1.00					
D2	0.22	0.15	0.24	0.00	1.00				
D3	-0.30	0.05	-0.07	0.19	-0.47	1.00			
P1	0.23	0.31	0.04	0.60	0.02	-0.07	1.00		
P2	-0.15	-0.33	0.00	-0.42	0.38	-0.19	-0.46	1.00	
P3	-0.05	0.05	-0.03	-0.12	-0.41	0.26	-0.44	-0.60	1.00

Fonte: Desenvolvido pelo autor

As Tabelas (19), (20) e (21) apresentam a matriz de correlação (coeficiente de correlação de Pearson) entre os parâmetros ajustados da

Função Lognormal Bimodal, para as reações R5, R4 e R8, respectivamente. Considerando um nível de significância de 5%, os pontos destacados indicam correlação entre os parâmetros.

Tabela 19 Matriz de Correlação entre os parâmetros ajustados da Função Lognormal Bimodal [Reação – R5].

	M1	M2	D1	D2	P1	P2
M1	1.000					
M2	0.508	1.000				
D1	-0.642	0.225	1.000			
D2	-0.123	-0.628	-0.379	1.000		
P1	0.409	0.678	0.142	-0.561	1.000	
P2	-0.409	-0.678	-0.142	0.561	-1.000	1.000

Fonte: Desenvolvido pelo autor

Tabela 20 Matriz de Correlação entre os parâmetros ajustados da Função Lognormal Bimodal [Reação – R4]

	M1	M2	D1	D2	P1	P2
M1	1.000					
M2	0.514	1.000				
D1	-0.166	0.669	1.000			
D2	-0.285	-0.780	-0.529	1.000		
P1	0.418	0.737	0.522	-0.782	1.000	
P2	-0.418	-0.737	-0.522	0.782	-1.000	1.000

Fonte: Desenvolvido pelo autor

Tabela 21 Matriz de Correlação entre os parâmetros ajustados da Função Lognormal Bimodal [Reação – R8]

	M1	M2	D1	D2	P1	P2
M1	1.000					
M2	0.611	1.000				
D1	0.189	0.930	1.000			
D2	-0.362	-0.732	-0.568	1.000		
P1	0.450	0.868	0.839	-0.618	1.000	
P2	-0.450	-0.868	-0.839	0.618	-1.000	1.000

Fonte: Desenvolvido pelo autor

APÊNDICE F – Teste de Kolmogorov-Smirnov

Testes de aderência servem para comparar as distribuições de probabilidades empíricas de uma variável com as distribuições de probabilidades teóricas estimadas pela função de distribuição em teste. Esta distribuição teórica pode ser uma distribuição de probabilidade clássica (normal, exponencial ou log-normal), ou proporções definidas especificamente para o problema (BARBETA *et. al* 2009).

Os testes de *Kolmogorov-Smirnov* e Qui-quadrado são amplamente utilizados (ASSIS *et. al* 1996), sendo o primeiro, normalmente, mais eficaz que o segundo. Já o teste de *Lilliefors* é mais eficaz que o teste de *Kolmogorov-Smirnov*, porém é específico para verificar a aderência dos dados amostrais de distribuições normais.

Salienta-se ainda que o teste de aderência Qui-quadrado apresenta limitações, visto que a frequência de uma classe não pode ser inferior a cinco e os dados são agrupados em classes perdendo informações, o que não ocorre no teste de Kolmogorov-Smirnov, que além de poder ser realizado com os dados agrupados, pode também ser realizado com os dados isoladamente (CAMPOS, 1983).

O teste de Kolmogorov-Smirnov é um teste que detecta diferenças em relação à tendência central, dispersão e simetria e de acordo com BARBETA *et. al* (2009) pode ser assim definido:

Seja $F(x)$ a função de distribuição acumulada, com parâmetros especificados, para a qual se quer verificar a aderência dos dados. As hipóteses são:

H_0 : os dados provêm de $F(x)$ [há aderência].

H_1 : os dados não provêm de $F(x)$ [não há aderência].

Sejam as distribuições de frequência acumuladas: a empírica, $S(x)$, e a teórica, $F(x)$. Para cada elemento da amostra, obtém-se a diferença absoluta entre essas duas distribuições. A estatística do teste é a diferença absoluta máxima (d_{\max}). O procedimento é assim descrito:

Defini-se $S(x)$ para cada valor x_i ($i=1, 2, \dots, n$) como:

$$S(x_i) = \frac{\text{número de valores} \leq x_i}{n}$$

onde n é o tamanho da amostra e x_i é um valor qualquer da amostra.

Obtem-se, para cada valor x_i ($i = 1, 2, \dots, n$), os valores teóricos $F(x_i)$, calculados pela função de distribuição acumulada $F(x)$, especificada em H_0 .

Verifica-se a discrepância entre $S(x)$ e $F(x)$ através das diferenças absolutas entre $F(x_i)$ e $S(x_i)$, e entre $F(x_i)$ e $S(x_{i-1})$, para $i = 1, 2, \dots, n$.

Calcula-se a estatística de teste (d_{\max}) em termos da amostra em análise:

$$d_{m\acute{a}x} = m\acute{a}x_i \{ |F(x_i) - S(x_i)|, |F(x_i) - S(x_{i-1})| \}$$

Uma vez identificada a distância máxima $d_{m\acute{a}x}$, compara-se o valor com um valor indicado na Tabela (22) para $d_{cr\acute{it}ico}$ de acordo com o nível de significância (α) e do tamanho n da amostra. Regra de decisão:

Se $d_{m\acute{a}x} < d_{cr\acute{it}ico}$ então aceita H_0

Se $d_{m\acute{a}x} \geq d_{cr\acute{it}ico}$ então rejeita H_0

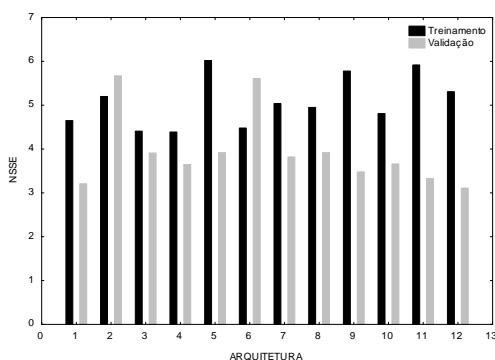
Tabela 22 Valores de $d_{cr\acute{it}ico}$ para o Teste de Kolmogorov-Smirnov com diferentes níveis de significância (α)

Tamanho da Amostra (N)	Nível de Significância (α)			
	0,20	0,10	0,05	0,01
1	0,900	0,950	0,975	0,995
2	0,684	0,776	0,842	0,929
3	0,565	0,642	0,708	0,829
4	0,494	0,564	0,624	0,734
5	0,446	0,510	0,563	0,669
6	0,410	0,470	0,521	0,618
7	0,381	0,438	0,486	0,577
8	0,358	0,411	0,457	0,543
9	0,339	0,388	0,432	0,514
10	0,322	0,368	0,409	0,486
15	0,268	0,304	0,338	0,404
20	0,231	0,264	0,294	0,352
25	0,210	0,240	0,264	0,320
30	0,190	0,220	0,242	0,290
40	-	-	0,210	0,250
50	-	-	0,190	0,230
N>50	$\frac{1,07}{\sqrt{N}}$	$\frac{1,22}{\sqrt{N}}$	$\frac{1,36}{\sqrt{N}}$	$\frac{1,63}{\sqrt{N}}$

APÊNDICE G – Determinação da Arquitetura da Rede Neural MLP

Nesse trabalho foram testadas arquiteturas de RNs MLP com uma e duas camadas ocultas, visto que segundo Cybenko, (1989) uma rede com uma camada intermediária pode qualquer função contínua e a utilização de duas camadas intermediárias permite a aproximação de qualquer função. Os testes foram realizados em triplicata, assim os resultados indicam a soma dos erros quadráticos normalizados [NSSE] juntamente com o desvio padrão na etapa de treinamento e validação. A Figura (38) mostra os resultados para o treinamento de RN com duas camadas ocultas em diferentes configurações, com o diâmetro médio como variável de saída.

Figura 38 Resultados da função objetivo [NSSE] para o treinamento e validação da rede neural MLP com diferentes arquiteturas



Fonte: Desenvolvido pelo autor

Os resultados da Tabela (23) indicam que entre as arquiteturas avaliadas, as arquiteturas codificadas como [1] e [12] se mostraram promissoras apresentando um número reduzido de neurônios em uma das camadas, sugerindo que a rede comporta-se melhor com uma camada predominante. Diante desses indícios, foram realizadas simulações onde há somente uma camada oculta na rede neural.

Tabela 23 Resultado da NSSE na inferência do diâmetro médio de partícula com diferentes arquiteturas para RNs MLP com duas camadas ocultas

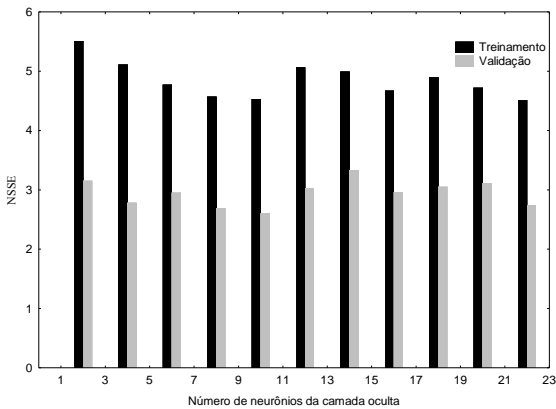
Arquitetura	1ª Camada Oculta	2ª Camada Oculta	NSSE Treinamento	NSSE Validação
1	18	2	4.65 ± 0.23	3.21 ± 0.17
2	16	12	5.20 ± 0.18	5.67 ± 0.67
3	16	10	4.41 ± 0.15	3.91 ± 0.41
4	16	4	4.39 ± 0.22	3.65 ± 0.39
5	14	6	6.02 ± 0.21	3.92 ± 0.51
6	12	16	4.48 ± 0.14	5.61 ± 0.53
7	12	8	5.05 ± 0.20	3.82 ± 0.45
8	10	10	4.95 ± 0.28	3.92 ± 0.36
9	8	16	5.78 ± 0.17	3.48 ± 0.41
10	8	12	4.81 ± 0.20	3.66 ± 0.44
11	6	14	5.92 ± 0.28	3.33 ± 0.47
12	4	16	5.32 ± 0.34	3.11 ± 0.54

Fonte: Desenvolvido pelo autor

A Figura (39) apresenta os resultados de treinamento e validação da rede neural com diferentes quantidades de neurônios em uma única camada oculta, com o diâmetro médio como variável da saída, onde se verifica indícios de que acima de 10 neurônios ocorre *over-fitting* já que os resultados para a validação apresentam involução a partir desta proposta de arquitetura.

Os resultados da Tabela (24) não indicam diferenças acentuadas, no entanto, com poucos neurônios na camada oculta tem-se resultados com maior desvio padrão, o que pode indicar certa instabilidade. Por fim, as simulações com um número grande de neurônios na camada oculta se mostraram igualmente precisas, mas com um maior custo computacional, perdendo, assim, eficiência.

Figura 39 Resultados da função objetivo [NSSE] para treinamento e validação da rede neural com diferentes números de neurônios na camada oculta na estimação do diâmetro médio de partícula.



Fonte: Desenvolvido pelo autor

Tabela 24 Resultado da NSSE e desvio padrão na inferência do diâmetro médio de partícula com diferentes números de neurônios na camada oculta.

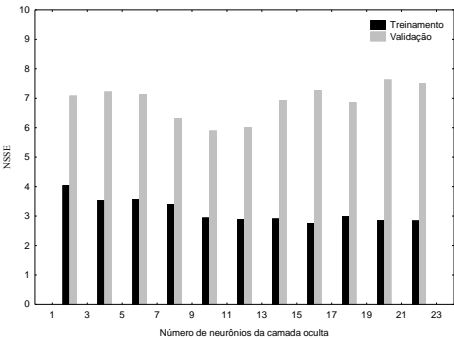
Número de Neurônios	NSSE	
	Treinamento	Validação
2	5,51 ± 0.87	3,15 ± 1.12
4	5,11 ± 0.46	2,78 ± 0.50
6	4,73 ± 0.35	2,95 ± 0.25
8	4,57 ± 0.14	2,69 ± 0.31
10	4,53 ± 0.16	2,60 ± 0.13
12	5,06 ± 0.17	3,03 ± 0.65
14	4,99 ± 0.15	3,33 ± 0.57
16	4,67 ± 0.16	2,96 ± 0.50
18	4,90 ± 0.34	3,05 ± 0.59
20	4,72 ± 0.36	3,11 ± 0.38
22	4,51 ± 0.23	2,74 ± 0.39

Fonte: Desenvolvido pelo autor

Como não há nenhuma garantia de que a arquitetura estabelecida até este momento seja coerente também para a estimação da conversão, já que os testes foram realizados com o diâmetro médio como variável de saída, é necessário realizar a otimização da arquitetura da rede neural admitindo a conversão como variável de saída.

A Figura (40) apresenta os resultados do treinamento e validação da rede neural MLP com uma única camada oculta e tendo como variável de saída à conversão. Nesse caso, o *over-fitting* não fica caracterizado, no entanto não se percebe melhora nos resultados de validação fazendo uso de uma arquitetura com mais que 10 neurônios na camada oculta, julgando assim desnecessário exceder este número.

Figura 40 Resultados do treinamento e validação da rede neural com diferentes números de neurônios na camada oculta na estimação da conversão



Fonte: Desenvolvido pelo autor

Os resultados que indicam o melhor desempenho da rede neural, composta por uma arquitetura com dez neurônios na camada oculta na estimação da conversão, estão apresentados na Tabela (25), e corroboram com os testes de arquiteturas que tinham como variável de saída o diâmetro médio de partícula. Vale ressaltar ainda que nessas condições o desvio padrão calculado é baixo, transmitindo maior consistência aos resultados.

Tabela 25 NSSE e desvio padrão na inferência da conversão com diferentes números de neurônios na camada oculta

Número de Neurônios	NSSE	
	Treinamento	Validação
2	4.04 ± 0,55	7.08 ± 0.82
4	3.54 ± 0.09	7.22 ± 0.49
6	3.57 ± 0.81	7.14 ± 0.84
8	3.40 ± 0.57	6.31 ± 0.38
10	2.94 ± 0.10	5.89 ± 0.33
12	2.89 ± 0.48	6.01 ± 0.44
14	2.91 ± 0.21	6.94 ± 0.57
16	2.75 ± 0.21	7.27 ± 0.29

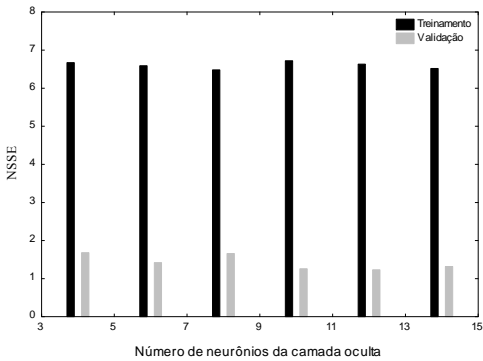
18	3.01 ± 0.23	6.86 ± 1.44
20	2.84 ± 0.36	7.63 ± 1.34
22	2.84 ± 0.29	7.51 ± 0.56

Fonte: Desenvolvido pelo autor

Desse modo, os testes preliminares indicaram que uma rede neural artificial MLP com uma única camada oculta composta por dez neurônios pode se mostrar eficiente na estimação da conversão e diâmetro médio de partícula.

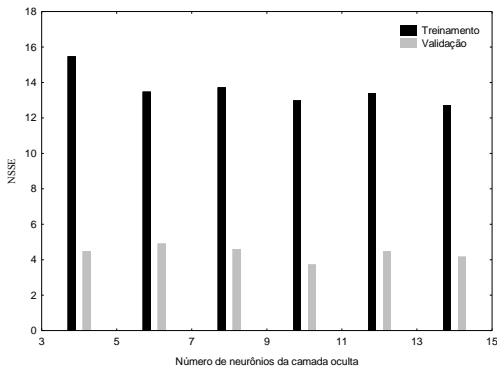
Esta mesma rede neural MLP é parte fundamental da proposta, visto que se busca a inferência por RNs da distribuição de tamanho de partícula. Como discutido no corpo do trabalho, a rede neural implementada realiza a inferência dos parâmetros da função lognormal bimodal que descrevem as DTPs. Por isso, é relevante realizar testes para determinar a arquitetura ideal da rede neural. Para tanto, os parâmetros M1 e D2 foram estimados usando diferentes arquiteturas, admitindo uma única camada oculta e variando o número de neurônios nesta camada. As Figuras (41) e (42) apresentam os resultados da função objetivo [NSSE] para o conjunto de treinamento e validação da rede neural na inferência do parâmetro M1 e D2, respectivamente, não sendo observadas diferenças significativas para diferentes números de neurônios na camada oculta.

Figura 41 Resultados do treinamento e validação da rede neural com diferentes números de neurônios na camada oculta na estimação do parâmetro M1



Fonte: Desenvolvido pelo autor

Figura 42 Resultados do treinamento e validação da rede neural com diferentes números de neurônios na camada oculta na estimação do parâmetro D2



Fonte: Desenvolvido pelo autor

Os resultados das Tabelas (26) e (27) corroboram com os resultados dos testes relativos a arquitetura da rede neural na inferência da conversão e diâmetro médio de partícula. Apesar de não haver diferença significativa, o treinamento da rede neural com 10 neurônios na camada oculta apresenta resultados satisfatórios com desvio padrão baixos, conferindo consistência à estratégia.

Tabela 26 NSSE e desvio padrão na inferência do parâmetro M1, com diferentes números de neurônios na camada oculta

Número de Neurônios	Erro Treinamento	Erro Validação
4	6.67 ± 0.33	1.68 ± 0.28
6	6.59 ± 0.08	1.42 ± 0.07
8	6.48 ± 0.01	1.66 ± 0.42
10	6.72 ± 0.13	1.26 ± 0.08
12	6.63 ± 0.12	1.23 ± 0.09
14	6.52 ± 0.13	1.32 ± 0.05
16	6.68 ± 0.37	1.42 ± 0.13

Fonte: Desenvolvido pelo autor

Tabela 27 NSSE e desvio padrão na inferência do parâmetro D2, com diferentes números de neurônios na camada oculta

Número de Neurônios	Erro Treinamento	Erro Validação
4	15.46 ± 0.05	4.49 ± 0.09
6	13.47 ± 0.10	4.89 ± 0.17
8	13.71 ± 0.16	4.60 ± 0.32
10	12.97 ± 0.13	3.71 ± 0.15
12	13.38 ± 0.45	4.48 ± 0.55
14	12.72 ± 0.66	4.16 ± 0.38
16	12.93 ± 0.48	5.16 ± 0.25

Fonte: Desenvolvido pelo autor

Por isso, fixar 10 neurônios na camada oculta para todos os casos investigados nesse trabalho não compromete a eficiência da rede neural, visto que a arquitetura é uma das mais efetivas. Além disso, facilita o processo de treinamento, pois não é necessário realizar readaptações quanto à arquitetura da rede para conjunto de dados de entrada e saída diferentes.